CYCLOTOMIC FIELDS AND FERMAT'S LAST THEOREM

TOM LOVERING

ABSTRACT. These are the course notes from the Harvard University spring 2015 tutorial on *Cyclotomic Fields and Fermat's Last Theorem*. Starting with Kummer's attempted proof of Fermat's Last Theorem, one is led to study the arithmetic of cyclotomic fields, in particular the p-part of the class group of $\mathbb{Q}(\zeta_p)$. Miraculously, it turns out these arithmetic questions can be answered by asking very simple questions about the Bernoulli numbers B_n, which can be defined and computed very explicitly. Our aim is to understand this connection, and exploit it to prove many cases of Fermat's Last Theorem.

CONTENTS

1. INTRODUCTION

Consider, for p a prime, the Fermat equation

$$x^p + y^p = z^p.$$

When $p = 2$, we can factorise it

$$x^2 = (z - y)(z + y)$$

and since we may assume x is odd and $gcd(y, z) = 1$, it follows that $z - y$ and $z + y$ are coprime, and hence there are (odd) integers a and b such that

$$z - y = a^2, z + y = b^2.$$

Rearranging, we see that

$$(x, y, z) = (ab, \frac{b^2 - a^2}{2}, \frac{b^2 + a^2}{2}).$$

Conversely, for any pair (a, b) of odd integers it is clear that

$$(ab)^2 + (\frac{b^2 - a^2}{2})^2 = \frac{4a^2b^2 + b^4 - 2a^2b^2 + a^4}{4} = (\frac{b^2 + a^2}{2})^2.$$

Thus we have solved the equation $x^2+y^2 = z^2$ completely (and there are lots of solutions).

It is reasonable to wonder if a similar method can be used to solve it in the case $p > 2$. Let us analyse the key steps.

(1) We needed to *factorise* the equation
$$x^2 = (z-y)(z+y).$$
For $x^p + y^p = z^p$ this doesn't look so useful: the only natural factorisations are things like
$$(x+y)(x^{p-1} - x^{p-2}y + ... + y^{p-1}) = z^p$$
and the second factor on the left hand side looks unmanagable.

However, suppose we enlarge our number system to include an element ζ such that $\zeta^p = 1$ (but $\zeta \neq 1$). Inside the complex numbers this is a standard procedure, and it can also be done abstractly. Then we are able to factorise the above further, and get
$$\prod_{k=0}^{p-1}(x + \zeta^k y) = z^p.$$

(2) We needed to show that each of the factors was coprime, and deduced the existence of a, b such that
$$z - y = a^2, z + y = b^2.$$
In the situation where p is odd, the factors live in the bigger number system $\mathbb{Z}[\zeta]$, so to make this argument we will need to figure out what concepts like 'coprime' mean and whether we can make the deduction that each factor is itself a pth power. It turns out these questions are subtle, and constitute the study of "the arithmetic of cyclotomic fields."

Given a number ring R (e.g. $R = \mathbb{Z}[\zeta]$) it may often not be the case that every number can be expressed uniquely as a product of primes, but this failure can be measured by a finite abelian group $Cl(R)$ called the *class group*. We will see that to make step (2) go through, we need a negative answer to the following.

Question 1.1. *Does p divide the order of $Cl(\mathbb{Z}[\zeta])$?*

How might one go about answering this question? For very small values of p, perhaps computing these class groups explicitly is practical, but even for $p > 23$ it becomes difficult for general methods to work.

2. "Proof" of First case of FLT

In this section we will use the ideas from the introduction to give a "proof" of FLT relying on a certain assumption. Analysing this assumption will be the task of much of the course.

Suppose $p > 3$ prime, and let ζ be a primitive pth root of unity. Let
$$\mathbb{Z}[\zeta] = \{a_0 + a_1\zeta + ... + a_{p-2}\zeta^{p-2} : a_i \in \mathbb{Z}\}$$
be the ring obtained by adjoining ζ to $\mathbb{Z}$ (we can add and multiply, but not generally divide). One is free to view this as a subring of $\mathbb{C}$, and it is clearly stable by complex conjugation.

If $I \subset \mathbb{Z}[\zeta]$ is an *ideal* of this ring, it makes sense to do arithmetic *modulo* I. This can be thought of either as arithmetic in the quotient ring $\mathbb{Z}[\zeta]/I$ or in $\mathbb{Z}[\zeta]$ itself with the equivalence relation that $\alpha \equiv \beta$ iff $\alpha - \beta \in I$. For example $p\mathbb{Z}[\zeta]$ is such an ideal, and we have the following result.

Lemma 2.1. *Let $\alpha \in \mathbb{Z}[\zeta]$. Then there always exists $a \in \mathbb{Z}$ such that*

$$\alpha^p \equiv a(\mod p).$$

Proof. Write out $\alpha = \sum_i a_i \zeta^i$. Then

$$\alpha^p \equiv \sum_i a_i^p \mod p$$

because $p|\binom{p}{r}$ for all $1 \leq r \leq p-1$. □

The *units* $\mathbb{Z}[\zeta]^*$ of this ring are those elements which have a multiplicative inverse. For example ζ is a unit because $\zeta\zeta^{p-1} = 1$, but 2 is not. We will need the following result about these units.

Lemma 2.2.

- *For any integers a, b not divisible by p,*
$$\frac{\zeta^a - 1}{\zeta^b - 1} \in \mathbb{Z}[\zeta]^*.$$
- *Any unit $\epsilon \in \mathbb{Z}[\zeta]^*$ can be expressed in the form*
$$\epsilon = \zeta^u \epsilon_0$$
where $\overline{\epsilon_0} = \epsilon_0$.

We postpone the proof of this lemma for now.

Lemma 2.3. *Let $\alpha = \sum_{i=0}^{p-1} a_i \zeta^i$ with $a_i \in \mathbb{Z}$, and suppose there is some i_0 such that $a_{i_0} = 0$.*

The if p divides α (in $\mathbb{Z}[\zeta]$), p divides each of the a_i (in $\mathbb{Z}$).

Proof. As a $\mathbb{Z}$-module, $\mathbb{Z}[\zeta]$ is freely generated by the elements $\{\zeta^i : 0 \leq i \leq p-1 : i \neq i_0\}$. In particular it is a basis for the $\mathbb{F}_p$-vector space $\mathbb{Z}[\zeta]/(p)$, from which the result follows. □

We will need to know that in $\mathbb{Z}[\zeta]$ the (rational) prime p is no longer prime: indeed it is a $p-1$st power, up to a unit.

Lemma 2.4. *We have*

$$(1-\zeta)^{p-1} = \epsilon p$$

for some unit ϵ.

Proof. Recall that the minimal polynomial of ζ is

$$X^{p-1} + X^{p-2} + ... + 1 = \prod_{i=1}^{p-1} (X - \zeta^i).$$

Substituting $X = 1$ and using that each $\frac{1-\zeta}{1-\zeta^i}$ is a unit, we recover the result. □

Finally, let us state the key assumption that will allow our proof to go through.

Assumption 2.5. *Suppose we have a product $\alpha_1...\alpha_k$ of elements in $\mathbb{Z}[\zeta]$ each of which is pairwise coprime (in the sense that the ideal (α_i, α_j) is the unit ideal), and that there is some $\beta \in \mathbb{Z}[\zeta]$ such that*

$$\alpha_1...\alpha_k = \beta^p.$$

Then each α_i can be writtten in the form

$$\alpha_i = \epsilon_i \beta_i^p$$

where ϵ_i is a unit and $\beta_i \in \mathbb{Z}[\zeta]$.

For example, if $\mathbb{Z}[\zeta]$ is a UFD, this assumption is clearly true. Any irreducible must divide the product a multiple of p times, but it only divides one of the factors by the pairwise coprimality assumption.

Theorem 2.6 (Conditional "first case" of Fermat's Last Theorem)**.** *Suppose the assumption is satisfied. Then there do not exist integers $x, y, z \in \mathbb{Z}$ such that $p \nmid xyz$ and*

$$x^p + y^p = z^p.$$

Proof. Suppose for contradiction that we have such a triple (x, y, z), and we may obviously assume x, y, z are pairwise coprime.

First a reduction: note that we may assume that

$$x \not\equiv y \mod p.$$

Indeed, if this cannot be realised by switching the variables x, y, z we must have

$$x \equiv y \equiv -z \mod p$$

but then $x^p + y^p - z^p \equiv 3x^p \not\equiv 0 \mod p$, a contradiction.

Now for the main argument. Inside $\mathbb{Z}[\zeta]$ we may factor

$$x^p + y^p = \prod_{i=0}^{p-1} (x + y\zeta^i).$$

We claim these each of factors are pairwise coprime. Indeed if some maximal ideal $\mathfrak{p}$ contains both $(x + y\zeta^i)$ and $(x + y\zeta^j)$ for $0 \leq i < j \leq p - 1$, then we also have

$$y(\zeta^i - \zeta^j) \in \mathfrak{p}.$$

Since $\mathfrak{p}$ is maximal, this implies at least one of y and $\zeta^i - \zeta^j$ is in $\mathfrak{p}$. If $y \in \mathfrak{p}$ then $x = (x+y\zeta^i) - \zeta^i y \in \mathfrak{p}$ also, which contradicts that x and y are coprime. But if $(\zeta^i - \zeta^j) \in \mathfrak{p}$ then $\mathfrak{p} = (1-\zeta)$ (which is visibly maximal since $\mathbb{Z}[\zeta]/(1-\zeta) \cong \mathbb{F}_p$). But if $(x+y\zeta^i) \subset (1-\zeta)$ this implies $(z) \subset (\prod_{i=1}^{p-1}(x + \zeta^i y)) \subset (1 - \zeta)^{p-1} = (p)$, contradicting that $p \nmid z$. Thus the coprimality claim is established.

Now we apply **??**, which tells us that each $x + \zeta^i y$ can be written in the form

$$x + \zeta^i y = \epsilon.\alpha^p$$

where ϵ is a unit and $\alpha \in \mathbb{Z}[\zeta]$. In particular, let us apply this where $i = 1$, and by **??** we can write

$$x + \zeta y = \zeta^u \epsilon_0 \alpha^p,$$

where $u \in \mathbb{Z}$, ϵ_0 is a real unit, and $\alpha \in \mathbb{Z}[\zeta]$.

We now go spoiling for a contradiction. By **??** we know that there is some $a \in \mathbb{Z}$ such that $\alpha^p \equiv a \mod p$. Now apply complex conjugation, and we get

$$x + \zeta^{-1} y \equiv \zeta^{-u} \epsilon_0 a \equiv \zeta^{-2u}(x + \zeta y) \mod p,$$

which we can rewrite as the relation

$$x + \zeta y \equiv \zeta^{2u} x + \zeta^{2u-1} y \mod p.$$

We now apply Lemma **??** many times. Firstly, if $1, \zeta, \zeta^{2u-1}, \zeta^{2u}$ are distinct, then (since $p \geq 5$) we get $p|x, y$ which clearly contradicts our assumptions. We can divide into three cases.

(1) Suppose $1 = \zeta^{2u}$, giving the equation

$$x + \zeta y \equiv x + \zeta^{-1} y$$

which reduces to $p|(\zeta^2 - 1)y$, which is only possible if $p|y$.

(2) Suppose $1 = \zeta^{2u-1}$, giving the equation

$$x + \zeta y \equiv \zeta x + y$$

which rearranges to $(x - y) - (x - y)\zeta \equiv 0$, implying by **??** that $p|x - y$. But we assumed $x \not\equiv y \mod p$.

(3) Suppose $\zeta = \zeta^{2u-1}$, giving the equation

$$x + \zeta y \equiv \zeta^2 x + \zeta y$$

which reduces to $p|(\zeta^2 - 1)x$ giving the same kind of contradiction as the first case.

And with contradictions in all cases, the theorem is proved. □

We must now investigate the validity of the assumption **??**. For example, are the $\mathbb{Z}[\zeta]$ always UFDs? Are they not but does the assumption hold nevertheless? Unfortunately in general the answer is "no," although for particular primes p it is often "yes" (whether it is for *infinitely many* p is an open problem).

The assumption $p \nmid xyz$ makes life considerably simpler, but can in fact be removed with a good amount of extra work, which we might do later in the course.

3. Galois Theory of Cyclotomic Fields

3.1. **Review of Galois Theory.**

3.1.1. Recall that a field K is a ring where every nonzero element is a unit. Equivalently the only ideals in K are (0) and (1).

3.1.2. The *characteristic* of a field K is the (non-negative generator of the) kernel of the canonical map $\mathbb{Z} \to K$. For example, the fields $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ all have characteristic 0 (they contain a canonical *copy* of $\mathbb{Z}$), whereas for p a prime, $\mathbb{Z}/p\mathbb{Z}$ is a field of characteristic $p > 0$. Unless stated otherwise, in this course all fields will have characteristic zero or be finite fields (which avoids our having to worry about *separability*).

3.1.3. Any ring homomorphism $f : K \to L$ between fields is automatically injective, and we call such a map a *field extension*. Such a map equips L with the structure of a K-vector space. If this is finite-dimensional, we say L/K is a *finite* extension. We define the *degree* of the extension to be

$$[L : K] := \dim_K L.$$

Given a chain $K \to L \to M$ of field extensions one always has the "Tower Law"

$$[M : L][L : K] = [M : K].$$

3.1.4. Let $f_1 : K \to L_1$ and $f_2 : K \to L_2$ be two field extensions. A K-*homomorphism* $g : L_1 \to L_2$ is a field homomorphism such that the two maps $g \circ f_1, f_2 : K \to L_2$ are equal. We write

$$Hom_K(L_1, L_2) := \{g : L_1 \to L_2 \text{ a } K\text{-homomorphism.}\}.$$

Lemma 3.2. *Suppose* $[L_1 : K]$ *is finite. Then*

$$|Hom_K(L_1, L_2)| \leq [L_1 : K].$$

Proof. Note that $[L_1 : K]$ is the dimension of the L_2-vector space $Hom_{K-Vec}(L_1, L_2)$, so if the lemma fails, there must be a dependence relation between the elements of $Hom_K(L_1, L_2)$ as a subset of this vector space. Let

$$\lambda_1\sigma_1 + ... + \lambda_k\sigma_k = 0$$

be the shortest such relation. Since the σ_i are distinct, we may find $x \in L_1$ such that $\sigma_1(x) \neq \sigma_k(x)$. But we see that for all $y \in L_1$,

$$0 = \sum_i \lambda_i\sigma_i(xy) = \sum_i \lambda_i\sigma_i(x)\sigma_i(y),$$

so the vector space map $\sum_i \lambda_i\sigma_i(x)\sigma_i$ is also identically zero. Subtracting $\sigma_k(x)$ times our original relation we obtain

$$\sum_i \lambda_i(\sigma_i(x) - \sigma_k(x))\sigma_i = 0$$

which is a nontrivial relation because $\sigma_1(x) \neq \sigma_k(x)$ but shorter than the one we started with because the kth term vanishes. □

3.2.1. We say that L_2 *splits* L_1 over K if equality holds in the previous lemma. We say a finite field extension $K \to L$ is *Galois* if it splits itself. In other words, L/K is Galois iff

$$|Aut_K(L)| = [L : K].$$

In this case, the finite group $Aut_K(L)$ is called the *Galois group of* L/K and often written $Gal(L/K)$.

Example 3.3. *As extensions of* $\mathbb{Q}$, $\mathbb{Q}(\sqrt{d}), \mathbb{Q}(\sqrt{d_1}, \sqrt{d_2}), \mathbb{Q}(\zeta_p)$ *are Galois, but* $\mathbb{Q}(\sqrt[3]{2})$ *is not.*

3.3.1. Here is another way to think about Galois extensions. Any field K (of characteristic 0) has an *algebraic closure* $\overline{K}$, which in general has a large group $Aut_K(\bar{K})$ of automorphisms. By the construction of $\bar{K}$, any finite extension of L admits (non-unique) K-homomorphisms $g : L \to \overline{K}$ (in fact L is split by $\bar{K}$). The extension L/K is Galois iff for any $\sigma \in \mathrm{Aut}_K(\bar{K})$ $\sigma(g(L)) \subset g(L)$ [prove it!].

The main reason Galois extensions are so popular is that their subextensions can be studied explicitly in terms of the finite group $Gal(L/K)$.

Theorem 3.4 (Fundamental theorem of Galois theory)**.** *Let* L/K *be a Galois extension. Then there is a natural bijection between:*

- *Subextensions* $K \to F \to L$.
- *Subgroups* $H \subset Gal(L/K)$.

This is given by

$$F \mapsto Gal(L/F) \subset Gal(L/K)$$

$$H \mapsto L^H = \{l \in L : \sigma(l) = l \forall \ \ l \in H\}.$$

If H *is normal in* $Gal(L/K)$ *then* L^H/K *is Galois with Galois group* $Gal(L/K)/H$*, and if* H *corresponds to* F *we always have*

$$|H| = [L : F].$$

For example, $Gal(\mathbb{Q}(\sqrt{2}, \sqrt{3})/\mathbb{Q}) \cong C_2 \times C_2$. This has three non-obvious subgroups of order 2, corresponding to the three quadratic subextensions $\mathbb{Q}(\sqrt{2}), \mathbb{Q}(\sqrt{3})$ and $\mathbb{Q}(\sqrt{6})$.

A Galois extension L/K is called *abelian* if $Gal(L/K)$ is an abelian group. In this case every subextension is Galois [why?].

Example 3.5. *Any degree two field extension* L/K *is Galois.*

3.6. **Galois theory of cyclotomic fields.** The main "Galois theoretic" result about cyclotomic fields is the following.

Theorem 3.7. *Let* $n \geq 3$*, and* ζ_n *a primitive* n*-th root of unity. Then* $\mathbb{Q}(\zeta_n)/\mathbb{Q}$ *is abelian, and the map*

$$\kappa : Gal(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \to (\mathbb{Z}/n\mathbb{Z})^*$$

given by taking σ *to the element* $i \in (\mathbb{Z}/n\mathbb{Z})^*$ *such that* $\sigma(\zeta_n) = \zeta_n^i$ *is an isomorphism of groups.*

Proof. One seems immediately that $\mathbb{Q}(\zeta_n)$ is Galois as the splitting field of $X^n - 1$. Since $\mathbb{Q}(\zeta_n)$ is generated by ζ_n over $\mathbb{Q}$, an automorphism σ is uniquely determined by where it sends ζ_n. This implies immediately that κ is injective, and it is easy to check it is a group homomorphism. The hard part is to prove surjectivity. Let $\Phi_n(X)$ be the nth cyclotomic polynomial (whose roots are the primitive nth roots of unity). Surjectivity of θ is equivalent to $\Phi_n(X)$ being irreducible (we are demanding that Galois acts transitively on the roots of $\Phi_n(X)$).

Suppose $\Phi_n(X)$ is reducible over $\mathbb{Q}$. Since $\Phi_n(X) \in \mathbb{Z}[X]$ we may assume by Gauss' lemma that $\Phi_n(X) = P(X)Q(X)$ where $P, Q \in \mathbb{Z}[X]$ and we assume P is irreducible. Let us assume $P(\zeta_n) = 0$ but k is such that $Q(\zeta_n^k) = 0$. By Dirichlet's theorem, there is some prime $p \equiv k \mod n$, so we know that $Q(\zeta_n^p) = 0$. This implies that

$$P(X)|Q(X^p).$$

Now, reduce mod p, and one gets $\bar{P}(X)|\bar{Q}(X^p) = \bar{Q}(X)^p$. In particular $\bar{P}$ and $\bar{Q}$ are not coprime in the UFD $\mathbb{F}_p[X]$. But $\overline{\Phi_n}(X)$ cannot have a repeated factor since $\frac{d(X^n-1)}{dX} = nX^{n-1}$ which is obviously coprime to $X^n - 1$ since $p \not| n$. This contradiction implies $\Phi_n(X)$ is irreducible, so the map is surjective, as required. □

Another key result which we will not prove (except perhaps at the end if people want to), but is very important to be aware of (for one's respect for cyclotomic fields) is the following.

Theorem 3.8 (Kronecker-Weber)**.** *Let $K/\mathbb{Q}$ be a finite abelian extension. Then there exists an n such that $K \subset \mathbb{Q}(\zeta_n)$. In other words, the cyclotomic fields contain all abelian extensions of $\mathbb{Q}$.*

For context, we note that these two theorems combined give what is called "class field theory" for the field $\mathbb{Q}$. They are the tip of a big and important iceberg.

4. Review of number fields

A *number field* is a finite extension K of $\mathbb{Q}$. For example, $\mathbb{Q}, \mathbb{Q}(\sqrt[5]{2})$ and $\mathbb{Q}(\zeta_n)$ are number fields, but $\bar{\mathbb{Q}}$ and $\mathbb{R}$ are not. In this chapter we will develop some basic tools for doing arithmetic in a number field in a way that generalises the usual arithmetic of $\mathbb{Z}$.

4.1. **Trace, norm and discriminant.** Let L/K be a finite extension of characteristic zero fields of degree d. Then viewing L as a K-vector space, any $l \in L$ can be viewed as a K-linear endomorphism $\times l : L \to L$. We define the *norm* $N_{L/K}(l)$ to be the determinant of this endomorphism, and the *trace* $tr_{L/K}(l)$ to be its trace. It is clear from basic linear algebra that

$$N_{L/K}(l_1 l_2) = N_{L/K}(l_1) N_{L/K}(l_2)$$

and

$$tr_{L/K}(l_1 + l_2) = tr_{L/K} l_1 + tr_{L/K} l_2.$$

Lemma 4.2. *Let $\tau_1, ..., \tau_d : L \hookrightarrow \bar{K}$ be the set of all K-embeddings of L in an algebraic closure. Then*

$$tr_{L/K}(l) = \tau_1(l) + \tau_2(l) + ... + \tau_d(l)$$

and

$$N_{L/K}(l) = \tau_1(l)\tau_2(l)...\tau_d(l).$$

Proof. Consider l as a $\bar{K}$-linear endomorphism of $L \otimes_K \bar{K} \cong \prod_{\tau_i} \bar{K}$, and with respect to the canonical basis on the right hand side the matrix of l is $Diag(\tau_1(l), ..., \tau_d(l))$. □

Now, suppose we are given a d-tuple of elements $\alpha_1, ..., \alpha_d \in L$. We define the *discriminant* to be

$$\Delta(\alpha_1, ..., \alpha_d) = det(tr_{L/K}(\alpha_i\alpha_j)).$$

Lemma 4.3. *The collection $\alpha_1, ..., \alpha_d$ is a K-basis for L if and only if $\Delta(\alpha_1, ..., \alpha_d) \neq 0$.*

Proof. Suppose they fail to be a basis, so there is a relation $\sum_i \lambda_i\alpha_i = 0$. Multiplying by α_j and taking the trace, we get

$$\sum_i \lambda_i tr_{L/K}(\alpha_i\alpha_j) = 0$$

for all j, which implies the matrix of $tr_{L/K}(\alpha_i\alpha_j)$ is singular, whence $\Delta = 0$.

Conversely, if $\Delta = 0$, fix a nontrivial solution (x_i) to the equations

$$\sum_i x_i tr_{L/K}(\alpha_i\alpha_j) = 0.$$

Then putting $\alpha = \sum_i (x_i\alpha_i)$, note that if (α_i) are a basis then we can express $\alpha^{-1} = \sum_j y_j\alpha_j$.

But then

$$tr_{L/K}(1) = tr_{L/K}(\alpha.\alpha^{-1}) = \sum_j y_j \sum_i x_i tr_{L/K}(\alpha_i\alpha_j) = 0$$

which is a contradiction because K has characteristic 0.[1] □

The following facts are useful for computing and manipulating discriminants. We leave their verification as an exercise.

Proposition 4.4. (1) *Let $\alpha_1, ..., \alpha_d$ and $\beta_1, ..., \beta_d$ be bases for L/K related by a change of basis matrix $\alpha_i = \sum a_{ij}\beta_j$. Then*

$$\Delta(\alpha_1, ..., \alpha_d) = \det(a_{ij})^2 \Delta(\beta_1, ..., \beta_d).$$

(2) *Using the embeddings $\tau_i : L \hookrightarrow \bar{K}$ we can compute*

$$\Delta(\alpha_1, ..., \alpha_d) = \det(\tau_i(\alpha_j))^2.$$

[1] In general one only needs the extension L/K to be separable, under which condition the trace pairing is always non-degenerate even if $tr(1) = 0$.

(3) *Suppose* $\beta \in L$ *such that* $1, \beta, .., \beta^{d-1}$ *are linearly independent over* K*, and let* f *be its minimal polynomial. Then*

$$\Delta(1, \beta, ..., \beta^{d-1}) = (-1)^{(n(n-1))/2} N_{L/K}(f'(\beta)).$$

4.5. **Rings of Integers in Number fields.** Let $K/\mathbb{Q}$ be a number field. An *algebraic integer in* K is a number $\alpha \in K$ which satisfies a monic polynomial

$$f(X) = X^d + a_{d-1}X^{d-1} + ... + a_1X + a_0$$

where each $a_i \in \mathbb{Z}$. We write $\mathcal{O}_K$ for the set of all such elements, which we call the *ring of integers of* K. This name is justified by the following theorem.

Theorem 4.6. *The set* $\mathcal{O}_K$ *is a subring of* K *containing* $\mathbb{Z}$*.*

Since $a \in \mathbb{Z}$ satisfies the monic polynomial $X - a = 0$, it is clear that $\mathbb{Z} \subset \mathcal{O}_K$. The rest of the theorem will proceed via some interesting commutative algebra lemmas.

Lemma 4.7 ("Cayley-Hamilton Theorem")**.** *Let* A *be a ring,* $I \subset A$ *an ideal, and* $M = (m_1, ..., m_k)$ *a finitely generated* A*-module. Whenever*

$$\phi : M \to M$$

is an endomorphism with $\phi(M) \subset I.M$ *then* ϕ *satisfies an equation (inside* $End_A(M)$*)*

$$\phi^k + a_{n-1}\phi^{k-1} + ... + a_0 = 0$$

where each $a_i \in I$*.*

Proof. Write $\phi(m_i) = \sum_j a_{ij}m_j$, where we may assume that each $a_{ij} \in I$. Then, working in the commutative ring $A[\phi] \subset End_A(M)$, the matrix $P_{ij} = \delta_{ij}\phi - a_{ij}$ kills every generator m_j of M. In particular $\det P.I_k = adj(P).P$ acts formally on any generator by

$$(\det P)(m_i) = \sum_j \delta_{ij} \det P(m_j) = \sum_{j,k} (adj(P)_{ik}P_{kj})(m_j) = \sum_k adj(P)_{ik}(\sum_j P_{kj}(m_j)) = 0.$$

Thus we have the relation

$$\det(\delta_{ij}\phi - a_{ij}) = 0$$

which can be expanded out to give a polynomial of the form required. □

Recall that if $A \subset B$ an inclusion of rings, we say $x \in B$ is *integral over* A if it satisfies a monic polynomial equation $b^n + a_{n-1}b^{n-1} + ... + a_0 = 0$ with coefficients $a_i \in A$.

Lemma 4.8 (Integrality lemma)**.** *If* $A \subset B$ *are rings, and* $x \in B$ *the following are equivalent.*

(1) x *is integral over* A*,*
(2) $A[x]$ *is finitely generated as an* A*-module,*
(3) $A[x] \subset C \subset B$ *where* C *is a subring finitely generated as an* A*-module,*
(4) *There exists a faithful* $A[x]$*-module* M*, finitely generated as an* A*-module.*

Proof. Firstly, (1) $\Rightarrow$ (2) is clear because if $x^n = a_{n-1}x^{n-1} + ... + a_0$ then $1, x, x^2, ..., x^{n-1}$ will do as a basis for $A[x]$ as an A-module. Next (2) $\Rightarrow$ (3) is clear taking $C = A[x]$ and (3) $\Rightarrow$ (4) is clear taking $M = C$.

The difficult part is (4) $\Rightarrow$ (1). But we may take $\times x : M \to M$ viewed as an A-module endomorphism and use the Cayley-Hamilton theorem with $I = A$ to recover (letting d be the number of generators needed to view M as a finitely generated A-module)

$$x^d + a_{d-1}x^{d-1} + ... + a_0 = 0$$

as an A-endomorphism of M. But M is faithful: i.e. $A[x] \to End_{A[x]}(M) \subset End_A(M)$ is injective, so if the above is the zero endomorphism it must be zero as an element of $A[x]$, proving that x is integral over A. □

Lemma 4.9 (Tower law (for rings))**.** *If $A \subset B \subset C$ a sequence of rings such that B is finitely generated as an A-module and C finitely generated as a B-module, then C is finitely generated as an A-module.*

Proof. If $B = Ax_1 + ... + Ax_n$ and $C = By_1 + ... + By_m$ then

$$C = Ax_1y_1 + Ax_2y_1 + ... + Ax_ny_m.$$

□

Proof. We are now in a position to prove that $\mathcal{O}_K$ is a ring. Suppose $x, y \in \mathcal{O}_K$. Then both are integral over $\mathbb{Z}$ in K, so by (1) $\Rightarrow$ (2) of the integrality lemma $\mathbb{Z}[x]$ is finitely generated over $\mathbb{Z}$ and $\mathbb{Z}[x, y]$ is finitely generated over $\mathbb{Z}[x]$, which implies by the tower law that $\mathbb{Z}[x, y]$ is finitely generated over $\mathbb{Z}$. In particular by applying (3) $\Rightarrow$ (1) of the integrality lemma to $\mathbb{Z}[x + y], \mathbb{Z}[xy] \subset \mathbb{Z}[x, y]$ we see that $x + y$ and xy are integral over $\mathbb{Z}$, so lie in $\mathcal{O}_K$. □

Example 4.10. *Let $K = \mathbb{Q}(\sqrt{2})$. The ring of integers consists of $\alpha = a + b\sqrt{2}$ satisfying a monic polynomial with $\mathbb{Z}$-coefficients. In this case we have*

$$\alpha^2 - Tr_{K/\mathbb{Q}}(\alpha)\alpha + N_{K/\mathbb{Q}}(\alpha) = 0$$

so $\alpha \in \mathcal{O}_K$ iff

$$Tr_{K/\mathbb{Q}}(\alpha) = 2a \in \mathbb{Z}$$

and

$$N_{K/\mathbb{Q}}(\alpha) = a^2 - 2b^2 \in \mathbb{Z}.$$

The first condition says $a = c/2$ for $c \in \mathbb{Z}$ but the second implies

$$c^2/4 - 2b^2 = d \in \mathbb{Z}.$$

Since $2b^2 = -c^2/4 - d$, $b = e/2$ is forced to be a half-integer also, but we then get the equation in $\mathbb{Z}$

$$c^2 - 2e^2 = 4d$$

which implies c is even and then e is even, so in fact we must have $a, b \in \mathbb{Z}$, and such a, b clearly give integral norms and traces. Thus

$$\mathcal{O}_K = \mathbb{Z}[\sqrt{2}] = \{a + b\sqrt{2} : a, b \in \mathbb{Z}\}.$$

Now we have a ring $\mathcal{O}_K$ which should behave inside K like $\mathbb{Z}$ does inside $\mathbb{Q}$. Let us study it further.

Lemma 4.11. (1) *We have* $\mathcal{O}_K \cap \mathbb{Q} = \mathbb{Z}$.
(2) *For* $\alpha \in \mathcal{O}_K$, $Tr_{K/\mathbb{Q}}(\alpha), N_{K/\mathbb{Q}}(\alpha) \in \mathbb{Z}$.

Proof. Suppose $p/q \in \mathbb{Q}$ (with p, q coprime) is integral over $\mathbb{Z}$. Then x satisfies a polynomial

$$x^n + a_{n-1}x^{n-1}.. + a_0 = 0$$

with $a_i \in \mathbb{Z}$. Clearing denominators we get

$$p^n + a_{n-1}qp^{n-1} + ... + q^n a_0 = 0$$

but q and p are coprime and yet q divides p^n, but this is only possible if $q = 1$.

For the statement about traces and norms, let F be a Galois extension of $\mathbb{Q}$ containing K, and recall that

$$Tr_{K/\mathbb{Q}}(\alpha) = \sum_{\tau: K \hookrightarrow F} \tau(\alpha).$$

But each $\tau(\alpha)$ will satisfy a monic polynomial with $\mathbb{Z}$-coefficients and so lie in $\mathcal{O}_F$. Thus $Tr_{K/\mathbb{Q}}(\alpha) \in \mathcal{O}_F \cap \mathbb{Q} = \mathbb{Z}$ by the first part of the lemma. Similarly for norm. □

Lemma 4.12. *Let* $x \in K$. *Then there is some* $a \in \mathbb{Z}$ *such that* $ax \in \mathcal{O}_K$.

Proof. Since $x \in K$ and K is a number field, x satisfies a polynomial with $\mathbb{Q}$-coefficients, and by clearing denominators we may assume

$$a_n x^n + a_{n-1}x^{n-1} + ... + a_0 = 0$$

with each $a_i \in \mathbb{Z}$.

But now

$$0 = a_n^{n-1}(a_n x^n + a_{n-1}x^{n-1} + ... + a_0) = (a_n x)^n + a_{n-1}(a_n x)^{n-1} + ... + a_0 a_n^{n-1}$$

witnesses that $a_n x \in \mathcal{O}_K$. □

Proposition 4.13. (1) *If* $I \subset \mathcal{O}_K$ *a nonzero ideal,* $I \cap \mathbb{Z}$ *is a nonzero ideal of* $\mathbb{Z}$.
(2) *Every nonzero ideal* $I \subset \mathcal{O}_K$ *contains a* $\mathbb{Q}$*-basis for* K.

Proof. Firstly we claim $I \cap \mathbb{Z} = m\mathbb{Z}$ for some $m \neq 0$. Since $\mathbb{Z}$ is a PID it suffices to prove $I \cap \mathbb{Z}$ is nonzero. But given $\alpha \in I$ we know it satisfies a polynomial

$$\alpha^n + a_{n-1}\alpha^{n-1} + ... + a_0$$

with $a_i \in \mathbb{Z}$, and conclude that $a_0 \in I$. We may assume $a_0 \neq 0$ dividing through by a power of α if necessary.

We conclude that $m\mathcal{O}_K \subset I$, so the first part is done, and for the second it will suffice to prove $\mathcal{O}_K$ contains a $\mathbb{Q}$-basis for K. But this follows from the previous lemma: take any $\mathbb{Q}$-basis for K and rescale by integers so that each element lies in $\mathcal{O}_K$ and it remains a $\mathbb{Q}$-basis. □

By (**??**) we know that if $K/\mathbb{Q}$ is a number field of degree d, and $\alpha_1, ..., \alpha_d \in \mathcal{O}_K$, the *discriminant*

$$\Delta(\alpha_1, ..., \alpha_d) = \det(Tr_{K/\mathbb{Q}}(\alpha_i \alpha_j)) \in \mathbb{Z}.$$

For any nonzero ideal $I \subset \mathcal{O}_K$ the previous lemma tells us we may find $\alpha_1, ..., \alpha_d$ with nonzero discriminant. We may conclude that there exist $\alpha_1, ..., \alpha_d \in I$ such that $|\Delta(\alpha_1, ..., \alpha_d)| \in \mathbb{N}$ is positive and minimal amongst all choices of d-tuple forming a $\mathbb{Q}$-basis for K.

Proposition 4.14. *Let $\alpha_1, ..., \alpha_d$ be elements of I making $|\Delta(\alpha_1, ..., \alpha_d)|$ minimal positive. Then I is generated by $\alpha_1, ..., \alpha_d$ as a free $\mathbb{Z}$-module.*

Proof. Suppose not, so there is some $\beta \in I$ with $\beta = \sum_i x_i \alpha_i$ and wlog $x_1 \in \mathbb{Q}$ not an integer. Write $x_1 = \theta + m$ where $\theta \in (0, 1)$, and consider the set $\beta_1 = \beta - m\alpha_1, \beta_i = \alpha_i (i \geq 2)$. Then the matrix of transition between the bases (α_i) and (β_i) is upper-triangular with determinant θ, so

$$\Delta(\beta_1, ..., \beta_d) = \theta^2 \Delta(\alpha_1, ..., \alpha_d)$$

contradicting minimality of $|\Delta(\alpha_1, ..., \alpha_d)|$.

That I is a free $\mathbb{Z}$-module follows from $\alpha_1, ..., \alpha_d$ being a basis over $\mathbb{Q}$, so any expression of an element of I as a $\mathbb{Z}$-linear combination is even unique amongst $\mathbb{Q}$-linear combinations. □

Given any two integral bases for I their transition matrix will have determinant ± 1 and so the discriminants of the two bases will be equal. We write $\Delta(I)$ for this value, and in particular obtain the fundamental invariant

$$\delta_K := \Delta(\mathcal{O}_K)$$

the *discriminant* of the number field K.

Example 4.15. *Let $K = \mathbb{Q}(i)$. One can check that the ring of integers is $\mathbb{Z}[i]$, with $\mathbb{Z}$-basis $1, i$. To compute the discriminant we note $Tr(1) = 2, Tr(i) = 0, Tr(i^2) = -2$, so*

$$\delta_{\mathbb{Q}(i)} = \Delta(1, i) = -4.$$

In particular, we see that the discriminant may be negative.

A crucial property of number rings is the following.

Corollary 4.16. *For any nonzero ideal $I \subset \mathcal{O}_K$, $\mathcal{O}_K/I$ is finite.*

Proof. We know $I \supset a\mathcal{O}_K$ for $a \in \mathbb{Z}$, so it suffices for $\mathcal{O}_K/a\mathcal{O}_K$ to be finite. But $\mathcal{O}_K \cong \mathbb{Z}^d$ as a $\mathbb{Z}$-module, so clearly $|\mathcal{O}_K/a\mathcal{O}_K| = a^d$. □

Corollary 4.17. *The ring $\mathcal{O}_K$ is Noetherian (every ideal is finitely generated), and every nonzero prime ideal of $\mathcal{O}_K$ is maximal.*

Proof. The first part is immediate from the proposition, and for the second we note that if $I \subset \mathcal{O}_K$ is prime, $\mathcal{O}_K/I$ is a domain but by the previous corollary it is also finite. Since every finite domain is a field,[2] we conclude that $\mathcal{O}_K/I$ is a field and so I is maximal. □

[2]To prove every finite domain D is a field, take $a \in D$ nonzero. Since it is a domain, multiplication by a is injective, and since D is finite multiplication by a is bijective, so there is some x with $ax = 1$, proving a has an inverse.

Proposition 4.18. *The ring $\mathcal{O}_K$ is integrally closed in K.*

Proof. The key is we now know that $\mathcal{O}_K$ is finitely generated over $\mathbb{Z}$. If x is integral over $\mathcal{O}_K$, then by the integrality lemma $\mathcal{O}_K[x]$ is finitely generated over $\mathcal{O}_K$ but by the tower law for rings we conclude that $\mathcal{O}_K[x]$ is finitely generated over $\mathbb{Z}$. Thus by (2) $\Rightarrow$ (1) in the integrality lemma we see that x is integral over $\mathbb{Z}$, so in fact $x \in \mathcal{O}_K$. □

4.19. **Finiteness of the class group.** We would like $\mathcal{O}_K$ to enjoy a property like that of $\mathbb{Z}$ whereby any number can be factored uniquely into primes. Unfortunately as stated this property does not hold. For example, the ring of integers of $\mathbb{Q}(\sqrt{-5})$ is $\mathbb{Z}[\sqrt{-5}]$ and we have the identity

$$6 = 2 \times 3 = (1 - \sqrt{-5})(1 + \sqrt{-5})$$

which witnesses a number that can be factored into primes in two distinct ways.

However it turns out that nevertheless one is able to prove two important theorems in this direction, to which we now turn. Firstly, we will define an invariant called the *class group* of K which measures the failure of K to have the unique factorisation property, and prove it is a finite abelian group. Secondly, we will establish that although numbers in $\mathcal{O}_K$ cannot be factored uniquely into products of primes, *ideals* still can.

We first establish enough about ideals to define the class group.

Lemma 4.20. (1) *Let $I \subset \mathcal{O}_K$ be a nonzero ideal. If $x \in K$ is such that $xI \subset I$ then $x \in \mathcal{O}_K$.*
(2) *If $I, J \subset \mathcal{O}_K$ are nonzero ideals and $I = IJ$ then $J = \mathcal{O}_K$.*
(3) *If $I, J \subset \mathcal{O}_K$ are nonzero ideals and $\alpha \in \mathcal{O}_K$ is such that $\alpha I = JI$ then $J = (\alpha)$.*

Proof. Note that (1) follows immediately from the Cayley-Hamilton theorem. For (2) we may take $\alpha_1, ..., \alpha_d$ an integral basis for I, and note that we may write $\alpha_i = \sum_j \beta_{ij}\alpha_j$ with $\beta_{ij} \in J$. It follows that the determinant of the matrix β_{ij} is ± 1, so $1 \in J$, as required.

Finally for (3), we see that for any $\beta \in J$, $\beta/\alpha I \in I$, which implies in particular by (1) that $\beta/\alpha \in \mathcal{O}_K$. Thus $J \subset (\alpha)$ so $J\alpha^{-1} \subset \mathcal{O}_K$ is an ideal. But by assumption $J\alpha^{-1}I = I$, and so by (2) we conclude that $J\alpha^{-1} = \mathcal{O}_K$. I.e. $J = (\alpha)$. □

We may now define the class group. Two ideals $I, J \subset \mathcal{O}_K$ are equivalent (write $I \sim J$) if there are $x, y \in \mathcal{O}_K$ such that

$$xI = yJ.$$

The set $Cl(K)$ of equivalence classes is called the *class group*. Transitivity is the only non-obvious part of the check that this forms an equivalence relation, and follows from the observation that if $xI = yJ$ and $aJ = bK$ then $xaI = yaJ = ybK$. We will postpone the proof that the set of classes has a group structure until we have shown it is finite.

We remark in passing that I is equivalent to $\mathcal{O}_K$ iff there are $x, y \in \mathcal{O}_K$ such that $xI = (y)$, but then $I = (y/x)$ is principal, and conversely it's obvious that any principal ideal is equivalent to $\mathcal{O}_K$. Thus $\mathcal{O}_K$ is a PID (every ideal is principal) iff $|Cl(K)| = 1$.

To establish finiteness, we need the following "geometry of numbers" lemma, which generalises the "division algorithm" from the arithmetic of $\mathbb{Z}$.

Lemma 4.21 (Hurwitz' Lemma)**.** *There exists a positive integer M depending only on K with the property that for any pair $x, y \in \mathcal{O}_K$ with $y \neq 0$, one can find an integer t with $1 \leq t \leq M$ and $z \in \mathcal{O}_K$ such that*

$$|N_{K/\mathbb{Q}}(tx - zy)| < |N_{K/\mathbb{Q}}(y)|.$$

Proof. Let $w = x/y \in K$. Then the problem becomes whether for any $w \in K$ we can choose $z \in \mathcal{O}_K$ and $1 \leq t \leq M$ such that $|N(tw - z)| < 1$. Fix a $\mathbb{Z}$-basis $e_1, ..., e_d$ for $\mathcal{O}_K$, and write

$$w = \sum_i \lambda_i e_i, z = \sum_i z_i e_i$$

with $\lambda_i \in \mathbb{Q}, z_i \in \mathbb{Z}$.

Viewing $t(\lambda_1, ..., \lambda_d) : 1 \leq t \leq M$ as elements of $[0, 1)^d$ by taking fractional parts $\lambda_i \mapsto \{\lambda_i\}$, and dividing this up into m^d little cubes for some $m < \sqrt[m]{M}$, by the pigeonhole principle we get two elements t_1, t_2 giving $\{t_1\lambda_i\}$ and $\{t_2\lambda_i\}$ in the same small cube, and so

$$0 \leq (t_2 - t_1)\{\lambda_i\} < 1/m \text{ or } 1 - 1/m \leq (t_2 - t_1)\lambda_i < 0.$$

Taking $t = t_2 - t_1$, letting z_i be chosen such that $\mu_i = t\lambda_i - z_i \in [-1/m, 1/m)$ we obtain the estimate

$$|N(tw - z)| = |N(\sum_i \mu_i e_i)| = |\prod_j (\sum_i \mu_i \tau_j(e_i))| \leq C(1/m)^d$$

where $C = \prod_j (\sum_i |\tau_j(e_i)|)$.

Thus if we take $m > \sqrt[d]{C}$, and $M = m^d + 1$ we get the bound needed. □

Theorem 4.22. *The class group $Cl(K)$ is finite.*

Proof. Note that the ideal $(M!)$ is contained in only finitely many ideals. We prove the proposition by establishing that any ideal is equivalent to one containing $(M!)$.

Indeed, let I be a nonzero ideal, and $x \in I$ a nonzero element with minimal $|N(x)|$. For any $y \in I$, the lemma tells us we can find t with $1 \leq t \leq M$ and $z \in \mathcal{O}_K$ such that $|N(ty - zx)| < |N(x)|$. By the assumption on x, we conclude that $ty = zx$, and since $y \in I$ was general, we conclude that

$$M!I \subset (x).$$

Taking $J = M!/xI$ we get an ideal with $(M!)I = (x)J$ in particular equivalent to I, but also since $x \in I$, $M!x \in xJ$ so $M! \in J$. □

It is customary to write $h_K := |Cl(K)|$.

Corollary 4.23. (1) *For any nonzero ideal $I \subset \mathcal{O}_K$, there is a positive integer k such that I^k is principal.*

(2) *The set $Cl(K)$ has an abelian group structure such that*

$$[I][J] = [IJ].$$

Proof. Consider $I, I^2, ..., I^{h_K+1}$. Some pair of these must be equivalent, say I^i and I^j with $j > i$. We claim that if $k = j - i$, I^k is principal.

Indeed, there are $x, y \in \mathcal{O}_K$ such that $xI^j = yI^i$, which implies $I^j = (y/x)I^i$ and by (**??** (1)) we see $w = y/x \in \mathcal{O}_K$ and by (**??** (3)) $I^k = (w)$ is principal as required.

For (2) we first check the multiplication structure is well-defined: if $aI = a'I'$ and $bJ = b'J'$ then

$$(ab)IJ = (aI)(bJ) = (a'I')(b'J') = (a'b')I'J'.$$

It is obviously associative and commutative with identity $[(1)]$. By part (1) we can define the inverse of $[I]$ to be $[I^{k-1}]$ where k is such that I^k is principal. If $k < k'$ are two such positive integers, then clearly $I^{k'-k}$ is also principal, so I^{k-1} and $I^{k'-1}$ are equivalent and this notion is well-defined.

□

Note that (2) implies we may always take $k = h_K$ in part (1).

4.24. **Unique factorisation of ideals.** With finiteness of the class group in our pocket, we are ready to start doing arithmetic with ideals.

Lemma 4.25 (Cancellation law)**.** *If I, J, K are nonzero ideals of $\mathcal{O}_K$ and $IJ = IK$ then $J = K$.*

Proof. Suppose $I^k = (x)$ is principal. We deduce that $xJ = xK$ which obviously implies $J = K$. □

Lemma 4.26 ("To contain is to divide")**.** *If I, J are ideals with $I \supset J$, then there exists another ideal K such that*

$$J = IK.$$

Proof. Suppose $I^k = (x)$ is principal. Then $(x) \supset I^{k-1}J$, so we can take $K = x^{-1}I^{k-1}J$ which clearly does the job. □

Now we have shown ideals behave in many ways like numbers, let us state the main theorem of this section.

Theorem 4.27 (Fundamental theorem of (higher) arithmetic)**.** *Every nonzero ideal $I \subset \mathcal{O}_K$ can be expressed uniquely (up to re-ordering) as a product*

$$I = \mathfrak{p}_1^{e_1}...\mathfrak{p}_r^{e_r}$$

of prime ideals.

We prove this as a consequence of two lemmas.

Lemma 4.28. *Every nonzero ideal I can be expressed as a product of prime ideals.*

Proof. Let us prove this by induction on the finite number $|\mathcal{O}_K/I|$ (noting as a base case that $I = \mathcal{O}_K$ is a product of the empty set of primes). Firstly we observe that I must be contained in a maximal ideal $\mathfrak{p}_1$ (either by a general result or using the fact that $\mathcal{O}_K/I$ is finite). To contain is to divide, so $I = \mathfrak{p}_1 I'$, and clearly $|\mathcal{O}_K/I'|$ is strictly smaller, so I' can be written as a product of primes by induction. □

For $\mathfrak{p}$ a prime ideal and I a nonzero ideal let us define $ord_{\mathfrak{p}}(I)$ to be the unique non-negative integer k such that $\mathfrak{p}^k \supset I$ but $\mathfrak{p}^{k+1} \not\supset I$.

Lemma 4.29. *Let $I, J \subset \mathcal{O}_K$ be nonzero ideals, and $\mathfrak{p}$ a prime ideal. Then:*

(1) $ord_{\mathfrak{p}}(\mathfrak{p}) = 1$,
(2) *For $\mathfrak{p}' \neq \mathfrak{p}$, $ord_{\mathfrak{p}}(\mathfrak{p}') = 0$.*
(3) *We have*

$$ord_{\mathfrak{p}}(IJ) = ord_{\mathfrak{p}}(I) + ord_{\mathfrak{p}}(J).$$

Proof. For (1) the only thing to check is $\mathfrak{p}^2 \neq \mathfrak{p}$, which is clear from the cancellation law. For (2) note that if $\mathfrak{p} \supset \mathfrak{p}'$ then $\mathfrak{p} = \mathfrak{p}'$ since all primes are maximal.

For (3) letting $ord_{\mathfrak{p}}(I) = k, ord_{\mathfrak{p}}(J) = l$ we may write $I = \mathfrak{p}^k I', J = \mathfrak{p}^l J'$ and clearly $\mathfrak{p}^{k+l} \supset IJ$. Since to contain is to divide we know that $\mathfrak{p} \not\supset I', J'$, so since $\mathfrak{p}$ is prime $\mathfrak{p} \not\supset I'J'$. Hence $\mathfrak{p}^{k+l+1} \not\supset IJ$ and we are done. □

Let us now prove the theorem. By the first lemma we may write

$$I = \mathfrak{p}_1^{e_1} ... \mathfrak{p}_r^{e_r}$$

in the form claimed (and we assume the $\mathfrak{p}_i$ are distinct). By the second lemma we see that $e_i = ord_{\mathfrak{p}_i}(I)$ is uniquely determined.

4.30. **Splitting behaviour of primes and ramification.** Let $\mathfrak{p}$ be a nonzero prime ideal of $\mathcal{O}_K$. Since $\mathcal{O}_K/\mathfrak{p}$ is a field, $\mathfrak{p}$ must contain a unique rational prime $p = char(\mathcal{O}_K/\mathfrak{p})$. We can read off two statistics attached to $\mathfrak{p}$.

- The *ramification index* of $\mathfrak{p}$ is

  $$e_{\mathfrak{p}} := ord_{\mathfrak{p}}(p).$$

  We say that p *ramifies* in $K/\mathbb{Q}$ if there is some $\mathfrak{p}$ dividing p with $e_{\mathfrak{p}} > 1$.
- The *inertia degree* $f_{\mathfrak{p}}$ of $\mathfrak{p}$ is defined by the equation

  $$|\mathcal{O}_K/\mathfrak{p}| = p^{f_{\mathfrak{p}}}.$$

  Another way of saying this is that it is the degree of the extension of finite fields

  $$f_{\mathfrak{p}} = [\mathcal{O}_K/\mathfrak{p} : \mathbb{Z}/p\mathbb{Z}].$$

Consider a rational prime p and let us ask the following question. What is the shape of the factorisation of $p\mathcal{O}_K$ into primes of $\mathcal{O}_K$? First a lemma.

Lemma 4.31 (Chinese Remainder Theorem)**.** *Let A be a ring and $I_1, I_2, ..., I_k$ ideals such that $I_i + I_j = A$ is the unit ideal for any pair i, j. Write $I = I_1 I_2 ... I_k$ for their product. Then there is a natural isomorphism of rings*

$$\theta : A/I \xrightarrow{\cong} A/I_1 \times A/I_2 \times ... \times A/I_k.$$

Proof. The natural isomorphism is just given by projection $A/I \twoheadrightarrow A/I_i$ onto each factor. If $\theta(a) = 0$ then $a \in I_i$ for all i.

It is a surjection: let us show $(1, 0, 0, ..., 0)$ can be hit. This is equivalent to saying there is some $a \in I_2 \cap I_3 \cap ... \cap I_k$ such that $a - 1 \in I_1$. But $(I_1 + I_2)(I_1 + I_3)....(I_1 + I_k) = A$, which can be simplified to

$$I_1 + I_2...I_k = A.$$

Taking $1 \in A$ and writing $1 = a_1 + a$ we get the element a required.

It is also an injection: clearly the kernel is $I_1 \cap I_2 \cap ... \cap I_k$, so we must show this is equal to I. But given $a \in I_1 \cap I_2 \cap ... \cap I_k$ we may use $I_1 + I_2....I_k = A$ to write $a = aa_1 + aa'$ with $a_1 \in I_1$ and $a' \in I_2...I_k$. But now we see that certainly $a \in I_1(I_2 \cap ... \cap I_k)$. Proceeding inductively gives the result.[3] □

Now let us apply this to

$$p\mathcal{O}_K = \mathfrak{p}_1^{e_{\mathfrak{p}_1}} \mathfrak{p}_2^{e_{\mathfrak{p}_2}} .....\mathfrak{p}_r^{e_{\mathfrak{p}_r}}.$$

The Chinese remainder theorem gives a ring isomorphism

$$\mathcal{O}_K/p\mathcal{O}_K \cong (\mathcal{O}_K/\mathfrak{p}_1^{e_{\mathfrak{p}_1}}) \times ... \times (\mathcal{O}_K/\mathfrak{p}_r^{e_{\mathfrak{p}_r}}).$$

We want to measure the size of each of these factors, so need the following lemma.

Lemma 4.32. *Let $\mathfrak{p}$ be a prime of $\mathcal{O}_K$ with inertial degree f. Then*

$$|\mathcal{O}_K/\mathfrak{p}^e| = p^{ef}.$$

Proof. The projection $\mathcal{O}_K/\mathfrak{p}^e \twoheadrightarrow \mathcal{O}_K/\mathfrak{p}^{e-1}$ has kernel $\mathfrak{p}^{e-1}/\mathfrak{p}^e$. If we can show this subgroup has order p^f we will be done by induction. By the cancellation law we can find $x \in \mathfrak{p}^{e-1}$ not in $\mathfrak{p}^e$. Moreover we know that $\mathfrak{p}^{e-1} = (x) + \mathfrak{p}^e$ because the factorisation of the right hand side can be nothing but the left. Thus the natural map $\mathcal{O}_K \to \mathfrak{p}^{e-1}/\mathfrak{p}^e$ given by $a \mapsto ax$ is surjective, and its kernel is equal to $\mathfrak{p}$, so induces an isomorphism of groups $\mathcal{O}_K/\mathfrak{p} \xrightarrow{\cong} \mathfrak{p}^{e-1}/\mathfrak{p}^e$. □

Putting everything together, we see that

$$p^{[K:\mathbb{Q}]} = |\mathcal{O}_K/p\mathcal{O}_K| = p^{e_{\mathfrak{p}_1} f_{\mathfrak{p}_1}} ....p^{e_{\mathfrak{p}_r} f_{\mathfrak{p}_r}}$$

from which we may conclude the following.

Proposition 4.33 (Ramification-Degree formula)**.** *Let K be a number field and p a rational prime with primes $\mathfrak{p}_1, ..., \mathfrak{p}_r$ in $\mathcal{O}_K$ dividing it. Then*

$$[K : \mathbb{Q}] = \sum_i e_{\mathfrak{p}_i} f_{\mathfrak{p}_i}.$$

[3] I guess it's rather easier to prove this for number rings using "to contain is to divide", but given how simple the proof of the general result is we felt it worth presenting.

Now, if $K/\mathbb{Q}$ is *Galois* one can say something stronger: the primes lying over p will all be conjugate to one another, and in particular one has $e_{\mathfrak{p}_1} = ... = e_{\mathfrak{p}_r} =: e_p$ and $f_{\mathfrak{p}_1} = ... = f_{\mathfrak{p}_r} = f_p$ and the formula simplifies to

$$[K : \mathbb{Q}] = re_pf_p.$$

In this case, if $G = Gal(K/\mathbb{Q})$ we define the *decomposition group of* $\mathfrak{p}$ to be

$$D_{\mathfrak{p}} = \{\sigma \in G | \sigma(\mathfrak{p}) = \mathfrak{p}\}.$$

This group acts naturally on the finite field $\mathcal{O}_K/\mathfrak{p}$, so we have a map which happens to be surjective

$$D_{\mathfrak{p}} \twoheadrightarrow Gal((\mathcal{O}_K/\mathfrak{p})/\mathbb{F}_p)$$

and its kernel is the *inertia group* $I_{\mathfrak{p}}$.

In terms of these groups, the above formula has the interpretation that $|D_{\mathfrak{p}}| = e_pf_p$ as the stabiliser of a transitive action on r primes, f_p is the order of $Gal((\mathcal{O}_K/\mathfrak{p})/\mathbb{F}_p)$ and e_p is the order of the inertia group at $\mathfrak{p}$.

Finally, we state an important result on which primes ramify.

Theorem 4.34. *Let K be a number field, and p a rational prime. Then p ramifies in K iff p divides the discriminant δ_K of K.*

In particular this implies that in any given K only finitely many primes ramify, and there is an important bound $|\delta_K| > 1$ telling us that if $K \neq \mathbb{Q}$ it must have some primes which ramify.

4.35. **Units in rings of integers.** To close this section we briefly review without proof the basic facts about units.

Firstly, recall that we write $\mathcal{O}_K^*$ for the set of *units* in $\mathcal{O}_K$: elements which have an inverse in $\mathcal{O}_K$. For example $\mathbb{Z}^* = \{\pm 1\}$.

Lemma 4.36. *We have that $\alpha \in \mathcal{O}_K$ is a unit iff $N_{K/\mathbb{Q}}(\alpha) = \pm 1$.*

Proof. If α is a unit, it has an inverse $\beta \in \mathcal{O}_K$ such that $\alpha\beta = 1$. But then

$$1 = N(1) = N(\alpha\beta) = N(\alpha)N(\beta)$$

so $N(\alpha), N(\beta)$, being in $\mathbb{Z}$, must be ± 1.

Conversely, work in a Galois closure $\tau : K \hookrightarrow F$, and recall that $N(\alpha) = \alpha.\prod_{\tau' \neq \tau} \tau'(\alpha)$. The second factor is an algebraic integer, and equal to $N(\alpha)/\alpha \in K$ so since $\mathcal{O}_K$ is integrally closed, $N(\alpha)/\alpha \in \mathcal{O}_K$. But if $N(\alpha) = \pm 1$ this implies α is a unit. □

Let us do two examples. Firstly, consider $\mathbb{Q}(\sqrt{-5})$. This has ring of integers $\mathbb{Z}[\sqrt{-5}]$, so a general element is of the form $\alpha = a + b\sqrt{-5}$. The above lemma tells us this is a unit iff

$$a^2 + 5b^2 = \pm 1.$$

Clearly the only solution is $b = 0, a = \pm 1$, so the only units are -1 and 1.

On the other hand, consider $\mathbb{Q}(\sqrt{7})$ with ring of integers $\mathbb{Z}[\sqrt{7}]$. The norm equation now becomes

$$a^2 - 7b^2 = \pm 1.$$

One can see the smallest positive solution to this equation is $a = 8, b = 3$, so one has a unit $8 + 3\sqrt{7}$, but it's not difficult to see that each power $(8 + 3\sqrt{7})^k : k \in \mathbb{Z}$ is distinct, and in fact a general unit of $\mathbb{Z}[\sqrt{7}]$ is of the form[4]

$$u = \pm(8 + 3\sqrt{7})^k : k \in \mathbb{Z}.$$

In the first case, we had abstractly the finite group $\mathbb{Z}[\sqrt{-5}]^* \cong C_2$, but in the second we have the infinite group $\mathbb{Z}[\sqrt{7}]^* \cong \mathbb{Z} \times C_2$. These fit into the framework of the following theorem (which we will not prove).

Theorem 4.37 (Dirichlet's Unit Theorem)**.** *Let K be a number field and suppose it has r_1 real embeddings and $2r_2$ complex embeddings. Let $\mu_K \subset \mathcal{O}_K^*$ be the group of roots of unity in $\mathcal{O}_K$. Then*

$$\mathcal{O}_K^* \cong \mathbb{Z}^{r_1+r_2-1} \times \mu_K.$$

A *basis of fundamental units* $\alpha_1, ..., \alpha_{r_1+r_2-1}$ is a collection of units which generate the unit group modulo μ_K.

In conclusion, we have shown that a ring of integers $\mathcal{O}_K$ of a number field has a lot of elegant properties, but also has associated two rather deep invariants: an ideal class group $Cl(K)$ and a group of units $\mathcal{O}_K^*$. Making a detailed study of these invariants even in very special cases often requires real work, as we will see in the case of cyclotomic fields.

5. Basic arithmetic of cyclotomic fields

We now specialise to the case of cyclotomic fields $K = \mathbb{Q}(\zeta_n)$ and some of their subfields. Recall we have already proved the irreducibility of cyclotomic polynomials, which in particular gives a natural isomorphism

$$Gal(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \xrightarrow{\cong} (\mathbb{Z}/n\mathbb{Z})^*.$$

After the last section, recall that the most basic questions about such fields are:

- What is the ring of integers $\mathcal{O}_K$?
- What is the discriminant δ_K? In particular, which primes ramify in $K/\mathbb{Q}$?

We begin with the special case where $n = p^k > 2$ is a prime power. Always let $\zeta = \zeta_n$, $K = \mathbb{Q}(\zeta_n)$.

Lemma 5.1. *Suppose $n = p^k$.*

(1) *The minimal polynomial of ζ (over $\mathbb{Q}$) is given explicitly by*

$$f(X) = \frac{X^{p^k} - 1}{X^{p^{k-1}} - 1} = X^{p^{k-1}(p-1)} + X^{p^{k-1}(p-2)} + ... + 1.$$

(2) *The norm*

$$N_{K/\mathbb{Q}}(1 - \zeta) = p.$$

[4]Note that $(8 + 3\sqrt{7})^{-1} = 8 - 3\sqrt{7}$.

(3) *For any pair* $a, b \in (\mathbb{Z}/p^k)^*$,

$$\frac{1-\zeta^a}{1-\zeta^b} \in \mathcal{O}_K^*.$$

(4) *The ideal* $(1-\zeta)$ *is prime in* $\mathbb{Z}[\zeta]$ *and*

$$(1-\zeta)^{p^{k-1}(p-1)} = (p).$$

In particular p *is totally ramified*[5] *in* $K/\mathbb{Q}$.

Proof. Part (1) is obvious (given irreducibility of cyclotomic polynomials). Part (2) follows from part (1) by substituting $X = 1$. For part (3) it will suffice by the norm criterion for units and part (2) to establish that $\frac{1-\zeta^a}{1-\zeta^b} \in \mathcal{O}_K$.

But take c such that $a \equiv bc \mod p^k$, and we get

$$\frac{1-\zeta^a}{1-\zeta^b} = \frac{1-\zeta^{bc}}{1-\zeta^b} = 1 + \zeta^b + ... + \zeta^{(c-1)b} \in \mathcal{O}_K.$$

Finally for (4), observe that

$$\mathbb{Z}[\zeta]/(1-\zeta) = \mathbb{Z}[X]/(f(X), 1-X) = \mathbb{Z}/(f(1)) = \mathbb{Z}/p\mathbb{Z}$$

and by (2) and (3)

$$(1-\zeta)^{p^{k-1}(p-1)} = (unit)N(1-\zeta) = (unit)p.$$

□

Proposition 5.2. *The ring of integers of* $\mathbb{Q}(\zeta_{p^k})$ *is* $\mathbb{Z}[\zeta_{p^k}]$, *and its discriminant is*

$$\Delta(\mathbb{Z}[\zeta_{p^k}]) = \pm p^{p^{k-1}(pk-k-1)}$$

where the sign is negative precisely if $p^k = 4$ *or* $p \equiv 3 \mod 4$.

Proof. We shall fix $\zeta = \zeta_{p^k}$ and begin by computing $\delta = \Delta(\mathbb{Z}[\zeta_{p^k}])$. Recall that this may be computed as

$$\Delta(1, \zeta, \zeta^2, ..., \zeta^{p^{k-1}(p-1)-1}) = \det(\zeta^{ij})^2.$$

where i varies between 0 and $p^{k-1}(p-1)-1$ and j varies over $(\mathbb{Z}/p^k\mathbb{Z})^*$.

But the matrix (ζ^{ij}) is a *Vandermonde matrix*[6] so we see that

$$\delta = \prod_{i>j} (\zeta^i - \zeta^j)^2$$

[5]We say a prime p is totally ramified in K if $(p) = \mathfrak{p}^d$ for some prime $\mathfrak{p}$ of $\mathcal{O}_K$ and where $d = [K : \mathbb{Q}]$.

[6]Given a degree d polynomial f with roots $\alpha_1, ..., \alpha_d$ in an algebraic closure, its Vandermonde matrix V_f has in the ith row the powers $1, \alpha_i, \alpha_i^2, ..., \alpha_i^{d-1}$. The point is that abstractly the determinant vanishes iff some $\alpha_i = \alpha_j$, so by the factor theorem

$$\det V_f = \pm \prod_{i>j} (\alpha_i - \alpha_j).$$

where i and j range over integers between 1 and p^k coprime to p. In particular, by (3) and (4) of the previous lemma, we see that

$$\delta = (unit) \prod_{i>j} p^{2w(i-j)}$$

where

$$w(i-j) = \frac{1}{p^{k-1}(p-1)} p^{v_p(i-j)}.$$

Thus the computation is reduced to the combinatorics of how many pairs $i \neq j$ of integers between 1 and p^k and prime to p are divisible by each power of p. Let us count the number of such pairs exactly twice by instead counting ordered pairs (i, j).

Fix i, for which there are $p^{k-1}(p-1)$ choices. The number of j such that $p \not| i-j$ is $p^{k-1}(p-2)$ and for $m \geq 1$, the number of j such that $p^m | i-j$ but p^{m+1} does not divide $i-j$ is exactly $p^{k-m-1}(p-1)$. Therefore

$$S := \sum_{(i,j):i\neq j} w(i-j) = \frac{1}{p^{k-1}(p-1)} p^{k-1}(p-1)(p^{k-1}(p-2) + \sum_{m=1}^{k-1} p^{k-m-1}(p-1)p^m)$$

which simplifies to

$$S = p^{k-1}(p-2+(p-1)(k-1)) = p^{k-1}(pk-k-1).$$

Finally note that

$$\delta = (unit) p^S$$

and since $\delta \in \mathbb{Z}$, we see that this gives the formula up to the sign ± 1.

One could obtain the sign by careful book-keeping, but we give the following alternative argument which gives a useful general fact.

Lemma 5.3. *Let F be any number field, and r_2 the number of pairs of complex embeddings. Then δ_F has sign $(-1)^{r_2}$.*

Proof. Fix $\alpha_1, ..., \alpha_d$ a $\mathbb{Z}$-basis for $\mathcal{O}_F$. As σ_j runs over all embeddings $F \hookrightarrow \mathbb{C}$, recall the formula

$$\Delta(\alpha_1, ..., \alpha_d) = (\det \sigma_j(\alpha_i))^2.$$

Consider the action of complex conjugation on the matrix $(\sigma_j(\alpha_i))$. Any real row will be fixed and any conjugate pair of complex rows will be swapped, so in particular

$$\overline{\det(\sigma_j(\alpha_i))} = (-1)^{r_2} \det(\sigma_j(\alpha_i)).$$

If r_2 is even, this tells us the determinant is real, so its square is positive. If r_2 is odd this tells us the determinant is pure imaginary, so its square is negative. □

Returning to our computation, we see that the sign will be negative iff r_2 is odd. Since $n > 2$ K admits no real embeddings, so $r_2 = [K : \mathbb{Q}]/2$ and this condition is equivalent to $4 \not| [K : \mathbb{Q}]$.

But for $p = 2, [K : \mathbb{Q}] = 2^{k-1}$ so the condition is equivalent to $k = 2$ (since $n = 2$ is excluded). For p odd,

$$[K : \mathbb{Q}] = p^{k-1}(p-1)$$

which is obviously indivisible by 4 iff $p \equiv 3 \mod 4$.

So we have pinned down the discriminant of $\mathbb{Z}[\zeta]$ exactly, and it remains to verify that this is the full ring of integers $\mathcal{O}_K$. We have that $\mathbb{Z}[\zeta] \subset \mathcal{O}_K$. Moreover we claim that $\mathcal{O}_K \subset \frac{1}{\delta}\mathbb{Z}[\zeta]$. Indeed, we have that $\delta = a^2\delta_K$ for some integer $a > 0$ which is the index of the $\mathbb{Z}$-submodule $\mathbb{Z}[\zeta] \subset \mathcal{O}_K$. But since a divides δ, in particular δ kills $\mathcal{O}_K/\mathbb{Z}[\zeta]$ which proves the claim.

Since we have shown δ is a power of p up to sign, if there is some $\alpha \in \mathcal{O}_K$ which fails to be in $\mathbb{Z}[\zeta]$ we may assume it lies in $p^{-1}\mathbb{Z}[\zeta]$ (after multiplying through by a power of p).

It therefore suffices, multiplying the whole situation by p, to check that

$$p\mathcal{O}_K \cap \mathbb{Z}[\zeta] = p\mathbb{Z}[\zeta].$$

Let us take as our basis $1, (1-\zeta), (1-\zeta)^2, ..., (1-\zeta)^{d-1}$, where it's convenient to set $d = p^{k-1}(p-1)$. Consider a general element

$$z = \sum_i a_i(1-\zeta)^i$$

where each $a_i \in \mathbb{Z}$, and we assume $z \in p\mathcal{O}_K$. We wish to prove each $a_i \in p\mathbb{Z}$. But this follows immediately by successively reducing moduli higher powers of $(1-\zeta)$ and subtracting once we can prove the following.

Lemma 5.4. *We have an equality of ideals in* $\mathbb{Z}$

$$(1-\zeta) \cap \mathbb{Z} = p\mathbb{Z}.$$

Proof. Since $(p) = (1-\zeta)^d$ it is clear one has $p \in (1-\zeta)$, but also $1 \notin (1-\zeta)$. Since $p\mathbb{Z}$ is a maximal ideal of $\mathbb{Z}$, the equality is obvious. □

This establishes that $p\mathcal{O}_K \cap \mathbb{Z}[\zeta] = p\mathbb{Z}[\zeta]$ which was all we needed to conclude $\mathcal{O}_K = \mathbb{Z}[\zeta]$. □

One consequence of calculating these discriminants is that we can immediately see which primes ramify. For example in $\mathbb{Q}(\zeta_{p^k})/\mathbb{Q}$ only the prime p ramifies (and the infinite prime if you count it). For general n, let us first prove a lemma about ramification in systems of field extension.

Lemma 5.5. (1) *Suppose $\mathbb{Q} \subset F \subset K$ are number fields, and p ramifies in $F/\mathbb{Q}$. Then p ramifies in $K/\mathbb{Q}$.*

(2) *Suppose $K = K_1....K_n$ is a number field written as a compositum of subfields and p fails to ramify in any $K_i/\mathbb{Q}$. Assume also that $K/\mathbb{Q}$ is abelian.*[7] *Then p fails to ramify in $K/\mathbb{Q}$.*

[7]This assumption is unnecessary.

Proof. For (1), let $p\mathcal{O}_F = \mathfrak{p}_1^{e_1}....\mathfrak{p}_r^{e_r}$ be the factorisation of (p) in the smaller number ring $\mathcal{O}_F$. If (wlog) $e_1 > 1$, then any $\mathfrak{q}$ in $\mathcal{O}_K$ dividing $\mathfrak{p}_1$ will occur with multiplicity greater than one in the factorisation of p in $\mathcal{O}_K$, so p ramifies in $K/\mathbb{Q}$.

For (2), use induction on n. Suppose p ramifies in $K/\mathbb{Q}$ and let $\mathfrak{p}$ be a prime of K such that $ord_{\mathfrak{p}}(p) > 1$. Since K is abelian, the decomposition and inertia groups $I_{\mathfrak{p}} \subset D_{\mathfrak{p}} \subset Gal(K/\mathbb{Q})$ are well-defined, and since $\mathfrak{p}$ is ramified in $K/\mathbb{Q}$, $I_{\mathfrak{p}}$ is nontrivial.

That K is the compositum of $K_1, ..., K_n$ tells us that the product of natural projections $Gal(K/\mathbb{Q}) \to \prod_i Gal(K_i/\mathbb{Q})$ is injective. Indeed if σ were in the kernel, since we can write any element $x \in K$ as an algebraic combination of elements of $K_1, ..., K_n$ we see that $\sigma(x) = x$, so $\sigma = 1$. In particular we see that $I_{\mathfrak{p}}$ has nontrivial image in $\prod_i Gal(K_i/\mathbb{Q})$ and so in $Gal(K_i/\mathbb{Q})$ for some i. This implies that the prime $\mathfrak{q}$ of K_i containing $\mathfrak{p}$ witnesses that p ramifies in $K_i/\mathbb{Q}$. □

Proposition 5.6. *Let $K = \mathbb{Q}(\zeta_n)$ be any cyclotomic field. Then p ramifies in $K/\mathbb{Q}$ iff $p|n$.*

Proof. If $p|n$, then p ramifies in $\mathbb{Q}(\zeta_p) \subset \mathbb{Q}(\zeta_n)$ so p ramifies in $\mathbb{Q}(\zeta_n)$ by part (1) of the lemma. Conversely, if p does not divide $n = q_1^{e_1}....q_r^{e_r}$ it fails to ramify in any $\mathbb{Q}(\zeta_{q_i^{e_i}})$ and so by part (2) of the lemma we have that p does not ramify in $\mathbb{Q}(\zeta_n)/\mathbb{Q}$ (since one sees easily that $\mathbb{Q}(\zeta_n)$ is a compositum of such fields). □

This allows us to read off the following, which is otherwise not so obvious without appealing to irreducibility of all cyclotomic polynomials (and indeed one can use this to give another proof).

Corollary 5.7. *If $gcd(m, n) = 1$, then $\mathbb{Q}(\zeta_n) \cap \mathbb{Q}(\zeta_m) = \mathbb{Q}$.*

Proof. Indeed, if not, the intersection is some $K \neq \mathbb{Q}$, but the set of primes which ramify is contained in those dividing both m and n, which is empty. But $\mathbb{Q}$ admits no extensions which are unramified everywhere. □

Another very nice thing we can do at this point is prove quadratic reciprocity by "pure thought."

Theorem 5.8 (Gauss' law of Quadratic Reciprocity)**.** *Given two distinct odd primes p, q, let $(p/q) = 1$ if p is a square mod q, $(p/q) = -1$ otherwise. Then*

$$(p/q)(q/p) = (-1)^{(p-1)(q-1)/4}.$$

Proof. Consider $\mathbb{Q}(\zeta_q)$ and note that q is the only prime which ramifies. Note that $Gal(\mathbb{Q}(\zeta_q)/\mathbb{Q}) \cong (\mathbb{Z}/q\mathbb{Z})^*$ has the index two subgroup $(\mathbb{Z}/q\mathbb{Z})^{*2}$ of squares mod q, which must fix a quadratic extension $\mathbb{Q}(\sqrt{q^*})$ of $\mathbb{Q}$. Since it can only be ramified at q, we see that $q^* = q$ if $q \equiv 1 \mod 4$, and $q^* = -q$ if $q \equiv 3 \mod 4$.

Now consider the automorphism $\sigma_p : \zeta \mapsto \zeta^p$ of $\mathbb{Q}(\zeta_q)$. It is clear that

$$\sigma_p(\sqrt{q^*}) = (p/q)\sqrt{q^*}.$$

On the other hand since p doesn't ramify in $\mathbb{Q}(\sqrt{q^*})$, σ_p acts trivially on $\mathbb{Q}(\sqrt{q^*})$ iff the polynomial $X^2 - q^*$ splits after reduction modulo p, which is equivalent to $(q^*/p) = 1$.

This analysis gives us that

$$(p/q)(q/p) = (q^*/p)(q/p) = ((-1)^{(q-1)/2}/p) = (-1)^{(q-1)(p-1)/4}$$

as required. □

We now need a general result which will make it easy to compute the ring of integers and discriminant of arbitrary cyclotomic fields.

Proposition 5.9. *Let K, F be two number fields which are linearly disjoint over $\mathbb{Q}$, $KF = K \otimes_{\mathbb{Q}} F$ their compositum, and suppose their discriminants are coprime. Then*

$$\mathcal{O}_{KF} = \mathcal{O}_K \mathcal{O}_F$$

and

$$\delta_{KF} = \delta_K^{[F:\mathbb{Q}]} \delta_F^{[K:\mathbb{Q}]}.$$

Proof. Firstly note that if M/L a finite extension of number fields, since $N_{M/L}(\mathcal{O}_M) \subset \mathcal{O}_L$, we are guaranteed that $\delta_L^{[M:L]} | \delta_M$. Thus since δ_F and δ_K are coprime we are guaranteed that $\delta_K^{[F:\mathbb{Q}]} \delta_F^{[K:\mathbb{Q}]} | \delta_{KF}$, so it suffices to find a basis for $\mathcal{O}_K \mathcal{O}_F$ giving this discriminant.

But if $\alpha_1, ..., \alpha_s$ a basis for $\mathcal{O}_K$ and $\beta_1, .., \beta_t$ for $\mathcal{O}_F$, we can compute the discriminant of the natural tensor basis $\alpha_i \beta_j$ as

$$\Delta(\alpha_i \beta_j) = \det(\tau_k(\alpha_i \beta_j))^2 = \Delta(\alpha_i)^{[F:\mathbb{Q}]} \Delta(\beta_j)^{[K:\mathbb{Q}]},$$

where the second equality is an elementary exercise in linear algebra. □

As a consequence, we get the following.

Theorem 5.10. *Let $n > 2$. Then the ring of integers of $K = \mathbb{Q}(\zeta_n)$ is*

$$\mathcal{O}_K = \mathbb{Z}[\zeta_n]$$

and the discriminant is

$$\delta_{\mathbb{Q}(\zeta_n)} = (-1)^{\phi(n)/2} \frac{n^{\phi(n)}}{\prod_{p|n} p^{\phi(n)/(p-1)}}.$$

Proof. Write

$$\mathbb{Q}(\zeta_n) = \mathbb{Q}(\zeta_{q_1^{e_1}}) ... \mathbb{Q}(\zeta_{q_r^{e_r}}).$$

Since each of these fields has discriminant only divisible by q_i and in particular pairwise coprime, and are also linearly disjoint by Corollary 5.7, the proposition gives us that

$$\mathcal{O}_K = \mathbb{Z}[\zeta_{q_1^{e_1}}] .... \mathbb{Z}[\zeta_{q_r^{e_r}}] = \mathbb{Z}[\zeta_n]$$

and we prove the discriminant formula by induction on the number of distinct prime factors.

Firstly if $n = p^k$ a prime power note that

$$(-1)^{\phi(p^k)/2} \frac{p^{k(p^{k-1}(p-1))}}{p^{p^{k-1}(p-1)/(p-1)}} = \pm p^{kp^{k-1}(p-1) - p^{k-1}} = \pm p^{p^{k-1}(pk-k-1)}$$

agreeing with our existing formula.

Now let $n = p^k n'$ where the formula is known for n' and $p \not| n'$. Then the previous proposition tells us that since $\phi(n) = \phi(n')\phi(p^k) = \phi(n')(p^{k-1}(p-1))$,

$$\begin{aligned}\delta_K &= ((-1)^{\phi(p^k)/2} p^{p^{k-1}(kp-k-1)})^{\phi(n')}((-1)^{\phi(n')/2}\frac{n'^{\phi(n')}}{\prod_{q|n'} q^{\phi(n')/(q-1)}})^{p^{k-1}(p-1)} \\ &= (-1)^{\phi(n)}\frac{n^{\phi}(n)}{\prod_{q|n} q^{\phi(n)/(q-1)}}.\end{aligned}$$

as required. □

6. Bernoulli Numbers and the Kummer Congruences

In this section we switch gears and introduce the *Bernoulli numbers* which have nothing obviously to do with cyclotomic fields.

You have probably all seen the formulae

$$1 + 2 + ... + (n-1) = \frac{n(n-1)}{2} = \frac{1}{2}n^2 - \frac{1}{2}n,$$

and

$$1^2 + 2^2 + ... + (n-1)^2 = \frac{n(n-1)(2n-1)}{6} = \frac{1}{3}n^3 - \frac{1}{2}n^2 + \frac{1}{6}n.$$

Bernoulli posed the question: can one always calculate polynomials in n for the expressions

$$S_m(n) = 1^m + 2^m + ... + (n-1)^m?$$

Here is a naive way to solve the problem. The binomial theorem tells us that

$$(k+1)^{m+1} - k^{m+1} = \sum_{i=0}^{m} \binom{m+1}{i} k^i.$$

Summing over k between 0 and $n-1$ we obtain

$$n^{m+1} = \sum_{i=0}^{m} \binom{m+1}{i} S_i(n).$$

This can be re-written as

$$S_m(n) = \frac{1}{m+1}(n^{m+1} - \sum_{i=0}^{m-1} \binom{m+1}{i} S_i(n)).$$

Thus we have an inductive formula for S_m in terms of the lower degree polynomials S_i. In particular, this tells us that each $S_m(n)$ is a polynomial of the form

$$S_m(n) = \frac{1}{m+1}n^{m+1} + B_{m,m}n^m + B_{m,(m-1)}n^{m-1} + ... + B_{m,1}n$$

where each $B_{i,j}$ is a rational number. We would like to compute these coefficients. After trying to do this for a while, it is natural to define an auxiliary sequence B_k recursively by $B_0 = 1$ and

$$(m+1)B_m = -\sum_{i=0}^{m-1} \binom{m+1}{k} B_k.$$

This is the sequence of *Bernoulli numbers* and one can easily compute them using the recursion, with the first few values being

$$B_0 = 1, B_1 = -\frac{1}{2}, B_2 = \frac{1}{6}, B_3 = 0, B_4 = \frac{-1}{30}, ...$$

This recursion can be encoded perhaps more conveniently (from a conceptual point of view) in the following way.

Lemma 6.1. *Let*

$$\frac{t}{e^t - 1} = \sum_{k=0}^{\infty} b_k \frac{t^k}{k!}$$

be the Taylor expansion of $f(t) = \frac{t}{e^t - 1}$ *about* $t = 0$. *Then for all* k,

$$b_k = B_k.$$

Proof. We have

$$t = (e^t - 1) \sum_{k=0}^{\infty} b_k \frac{t^k}{k!} = \sum_{i=1}^{\infty} \sum_{k=0}^{\infty} b_k \frac{t^{i+k}}{i!k!}$$

Comparing coefficients, t gives us $b_0 = 1$, and for $r \geq 2$, the t^r-coefficient gives us

$$0 = \sum_{k=0}^{r-1} b_k \frac{1}{(r-k)!k!}.$$

Multiplying through by $r!$ one recovers the recursion formula defining the Bernoulli numbers which proves inductively that $b_k = B_k$ for all k. □

With this fact established, it's easy to directly establish a formula for $S_m(n)$ in terms of Bernoulli numbers.

Proposition 6.2. *We may write*

$$S_m(n) = \frac{1}{m+1} \sum_{k=0}^{m} \binom{m+1}{k} B_k n^{m+1-k}.$$

Proof. We prove it by comparing Taylor expansions. First note that since $e^{kt} = \sum_i t^i / i!$, we have

$$1 + e^t + e^{2t} + ... + e^{(n-1)t} = \sum_{m=0}^{\infty} S_m(n) \frac{t^m}{m!}.$$

On the other hand we may write

$$1 + e^t + e^{2t} + ... + e^{(n-1)t} = \frac{e^{nt} - 1}{e^t - 1} = \frac{e^{nt} - 1}{t} \frac{t}{e^t - 1}.$$

The right hand side can be expanded as

$$\frac{e^{nt} - 1}{t} \frac{t}{e^t - 1} = \sum_{k=1}^{\infty} \sum_{j=0}^{\infty} n^k \frac{t^{k-1}}{k!} B_j \frac{t^j}{j!}.$$

Comparing coefficients after multiplying by $m!$ gives the result. □

We have so far defined our sequence of Bernoulli numbers in two different ways (once by recursion which is useful for computing them and once by generating function which is useful in proofs), and noted one useful property relating them to sums of positive powers of integers. We now turn to sums of *negative* powers. Recall that for $Re(s) > 1$ we can define the Riemann *zeta function* by

$$\zeta(s) = 1 + \frac{1}{2^s} + \frac{1}{3^s} + ...$$

and the sum converges absolutely so the definition makes sense.

An important classical problem is the evaluation of this function when s is an integer. Famously Euler proved that

$$\zeta(2) = 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + ... = \frac{\pi^2}{6}.$$

In fact Euler proved the following more general result, giving explicit formulas for s any even positive integer in terms of Bernoulli numbers.

Theorem 6.3 (Bernoulli numbers as special values of $\zeta(s)$)**.** *For any $m \geq 1$, we have that*

$$\zeta(2m) = \frac{(-1)^{m+1}}{2} \frac{(2\pi)^{2m}}{(2m)!} B_{2m}.$$

Proof. Following Euler, we will prove it using the "infinite partial fraction" expansion of the cotangent function

$$\cot x = \frac{1}{x} - 2 \sum_{n=1}^{\infty} \frac{x}{n^2\pi^2 - x^2}.$$

Let us multiply through by x and use a geometric series expansion to get

$$x \cot x = 1 - 2 \sum_{k=1}^{\infty} \zeta(2k) \frac{x^{2k}}{\pi^{2k}}.$$

Now re-write

$$x \cot x = ix \frac{e^{ix} + e^{-ix}}{e^{ix} - e^{-ix}} = ix + 2i \frac{x}{e^{2ix} - 1} = 1 + \sum_{n=2}^{\infty} B_n \frac{(2ix)^n}{n!}.$$

Comparing coefficients yields the result. □

As well as seeing their utility in describing other natural objects, we are slowly learning more about the Bernoulli numbers as a sequence. Let us record the following, which follow easily from what we know already.

Corollary 6.4. *We have the following facts about the Bernoulli numbers B_n.*

(1) *If $n > 1$ is odd, $B_n = 0$.*
(2) *The sign of B_{2m} is $(-1)^{m+1}$.*
(3) *$|B_{2m}/2m| \to \infty$ as $m \to \infty$.*

Proof. We note that (1) follows either because one checks easily that $\frac{t}{e^t-1} + \frac{t}{2}$ is an even function, or by the comparison of coefficients at the end of the preceding proof. We can see (2) from the previous theorem because $\zeta(2m) > 0$ when $m \geq 1$.

For (3) note that by the previous theorem we can do rather stronger. Indeed clearly $\zeta(2m) > 1$ for all $m \geq 1$, so

$$|B_{2m}| > \frac{2(2m)!}{(2\pi)^{2m}}.$$

This is stronger than the bound we need. Indeed, we can use the trivial estimate

$$(2m-1)! > 7^{2m-8}$$

to see

$$\frac{|B_{2m}|}{2m} > \frac{2}{7^8}(\frac{7}{2\pi})^{2m} \to \infty.$$

□

Given the Bernoulli numbers are a sequence of rational numbers, it is natural to wonder how complicated the denominators can become. The following theorem answers this question.

Theorem 6.5 (Clausen-von Staudt)**.** *For each $m \geq 1$ there is an integer $A_{2m} \in \mathbb{Z}$ such that*

$$B_{2m} = A_{2m} - \sum_{p-1|2m} 1/p.$$

Before we are able to prove this, we will need some intermediate congruences. Recall that $\mathbb{Z}_{(p)} = \{a/b \in \mathbb{Q} | a, b \in \mathbb{Z}, p \not| b\}$ is the subring of *p-integral* rational numbers, over which it makes sense to consider congruences modulo powers of p.

Lemma 6.6. *Suppose $m \geq 1$ and p prime. Then $pB_m \in \mathbb{Z}_{(p)}$. If m is even, then moreover we have*

$$pB_m \equiv S_m(p) \mod p.$$

Proof. Let us use the familiar binomial coefficient identity

$$\binom{m+1}{k} = \frac{m+1}{m-k+1}\binom{m}{k}$$

to rewrite

$$\begin{aligned} S_m(n) &= \frac{1}{m+1}\sum_{k=0}^{m}\binom{m+1}{k}B_k n^{m+1-k} \\ &= \sum_{k=0}^{m}\frac{1}{m-k+1}\binom{m}{k}B_k n^{m+1-k} \\ &= \sum_{k=0}^{m}\binom{m}{k}\frac{B_{m-k}n^{k+1}}{k+1}. \end{aligned}$$

We use this to prove the first part by induction. Clearly $pB_1 = -p/2 \in \mathbb{Z}_{(p)}$. Now assume $pB_1, ..., pB_{m-1}$ are p-integral, and put $n = p$ in the above identity to get

$$pB_m = S_m(p) - \sum_{k=1}^{m}\binom{m}{k}pB_{m-k}\frac{p^k}{k+1}.$$

The p-integrality of pB_m follows provided $p^k/(k+1)$ is always p-integral for $k \geq 1$, which is obvious.

For the second part, it will suffice to check that for each k, $p|\binom{m}{k}pB_{m-k}\frac{p^k}{k+1}$. Since m is even, for $k = 1$, $B_{m-1} = 0$ so there is nothing to check. For $k \geq 2$, it suffices to check $p^k/(k+1)$ is always divisible by p. Again this is clear because $k + 1 < 2^k \leq p^k$ for all $k \geq 2$. □

We are almost home: it remains to compute congruences for $S_m(p)$.

Lemma 6.7. *If $p - 1 \not| m$, $S_m(p) \equiv 0 \mod p$. If $p - 1|m$, $S_m(p) \equiv -1 \mod p$.*

Proof. Let g be a primitive root mod p (i.e. an element of $(\mathbb{Z}/p\mathbb{Z})^*$ which generates it as a cyclic group). Then

$$S_m(p) = 1^m + ... + (p-1)^m \equiv 1 + g^m + g^{2m} + ... + g^{(p-2)m}.$$

Thus

$$(g^m - 1)S_m(p) \equiv g^{(p-1)m} - 1 \equiv 0 \mod p.$$

If $p - 1 \not| m$ then $g^m \not\equiv 1$, so we deduce $S_m(p) \equiv 0$. On the other hand if $p - 1|m, g^{im} = 1$ for all i, so

$$S_m(p) = 1 + 1 + ... + 1 \equiv -1 \mod p.$$

□

Now we may finish the proof of the Clausen-von Staudt theorem. By the lemmas, we see that

$$A_{2m} = B_{2m} + \sum_{p-1|2m}\frac{1}{p}$$

is p-integral for all p. Indeed, if $p - 1 \not| 2m$,

$$pA_{2m} \equiv pB_{2m} \equiv S_{2m}(p) \equiv 0 \mod p,$$

and if $p-1|2m$,

$$pA_{2m} \equiv pB_{2m} + 1 \equiv S_{2m}(p) + 1 \equiv 0 \mod p.$$

But this implies A_{2m} is an integer, proving the theorem.

One might expect it is possible to find better congruences than those of our lemma. Indeed, without any real new ideas, one can strengthen to the following. Write $B_m = U_m/V_m$ with $U_m > 0$ and $gcd(U_m, V_m) = 1$.

Proposition 6.8. *if $m \geq 2$ is even, then for all $n \geq 1$,*

$$V_m S_m(n) \equiv U_m n \mod n^2.$$

In particular we note that if $p-1 \not| m$, this implies $S_m(p) \equiv B_m p \mod p^2$, which is stronger than our result above.

Proof. Again this rests on the identity

$$S_m(n) = \sum_{k=0}^{m} \binom{m}{k} \frac{B_{m-k} n^{k+1}}{k+1}.$$

Note first that if $p|n$, $k \geq 1$ and $p \neq 2, 3$ then $\binom{m}{k} \frac{B_{m-k} n^{k-1}}{k+1}$ is p-integral. Indeed this follows because by Clausen-von Staudt $v_p(B_{m-k}) \geq -1$, and that $v_p(n^{k-1}) \geq (k-1)$. First note that if $k = 1$ $B_{m-k} = 0$ so the result is obvious. For $k \geq 2$ it suffices for $v_p(k+1) \leq k-2$, which is obvious because $p \geq 5$ and $k+1 \leq 5^{k-2}$ for all $k \geq 3$, and for $k = 2$ it is clear.

We next show that at $p = 2, 3$ at least $p\binom{m}{k} \frac{B_{m-k} n^{k-1}}{k+1}$ is p-integral.

For $p = 2$, again if $k = 1$ we get zero for $m > 2$ and it's easy to see for $m = 2$. For $k > 1$ note that $B_{m-k} = 0$ unless k even or equal to $m-1$. If k is even, $k+1$ is odd, so the result is clear. If $k = m-1$ the number is $-n^{m-2}$ which is in particular p-integral.

For $p = 3$, $3|n$, if $k \geq 2$ then we know $k+1 \leq 3^{k-1}$ which gives the result needed, and again the $k = 1$ case is no concern.

Putting these estimates together, we see that the greatest common divisor d of n and the denominator of

$$\frac{S_m(n) - B_m n}{n^2}$$

divides 6. But Clausen-von Staudt implies that $6|V_m$ always, so multiplying through by V_m we get that

$$V_m S_m(n) - U_m n \equiv 0 \mod n^2$$

as was required. □

With these slightly stronger congruences in our pocket, we can now establish the following famous formula of Voronoi. For $x \in \mathbb{R}$, we use the notation $[x]$ to denote the *floor* of x: the unique integer k such that $k \leq x < k+1$.

Proposition 6.9 (Voronoi formula)**.** *Let n be a positive integer, $m \geq 2$ even, and $a > 0$ coprime to n. Then*

$$(a^m - 1)U_m \equiv ma^{m-1}V_m \sum_{j=1}^{n-1} j^{m-1}[ja/n] \mod n.$$

Proof. Write $ja = q_j n + r_j$ with $0 \leq r_j < n$. Of course $[ja/n] = q_j$ and since a, n are coprime, the r_j cover all nonzero residue classes mod n as $j = 1, ..., n-1$.

Note that

$$j^m a^m = (q_j n + r_j)^m \equiv r_j^m + mq_j n r_j^{m-1} \equiv r_j^m + m[ja/n]na^{m-1}j^{m-1} \mod n^2.$$

Summing over $j = 1, ..., n-1$ we obtain

$$a^m S_m(n) \equiv S_m(n) + mna^{m-1} \sum_{j=1}^{n-1} j^{m-1}[ja/n] \mod n^2.$$

By the previous proposition, multiplying both sides by V_m gives

$$(a^m - 1)U_m n \equiv V_m mna^{m-1} \sum_{j=1}^{n-1} j^{m-1}[ja/n] \mod n^2.$$

Dividing through by n gives the formula desired. □

We note the following simple consequence which gives our first information about *numerators* of Bernoulli numbers.

Corollary 6.10. *If $p - 1 \nmid m$, then $B_m/m \in \mathbb{Z}_{(p)}$.*

Proof. We know $B_m \in \mathbb{Z}_{(p)}$ by Clausen-von Staudt. Let $m = p^t m_0$. Obviously if $t = 0$ there is nothing to prove. If $t > 0$, put $n = p^t$ in the Voronoi formula, and we see

$$(a^m - 1)U_m \equiv 0 \mod p^t,$$

and since we are free to choose a of order divisible by $(p-1)$ we deduce that $p^t | U_m$ as required. □

We are now able to prove perhaps the most important congruences between Bernoulli numbers: the so-called *Kummer congruences* (although he is only responsible for the case $e = 1$).

Theorem 6.11 (First Kummer Congruences)**.** *Let $m, m' \geq 2$ be even, p a prime with $p - 1 \nmid m$. Then whenever*

$$m \equiv m' \mod p^{e-1}(p-1)$$

we have the congruence

$$(1 - p^{m-1})\frac{B_m}{m} \equiv (1 - p^{m'-1})\frac{B_{m'}}{m'} \mod p^e.$$

Proof. Let's warm up with the case $e = 1$, where we need to show that

$$\frac{B_m}{m} \equiv \frac{B_{m'}}{m'} \mod p.$$

Suppose $t = ord_p(m)$, so the previous corollary implies $p^t | U_m$. Taking $n = p^{t+1}$ in Voronoi we get

$$m^{-1}(a^m - 1)B_m \equiv a^{m-1} \sum_{j=1}^{p^{t+1}-1} j^{m-1}[ja/p^{t+1}] \mod p.$$

Since the terms where $p|j$ vanish, the right hand side is visibly unchanged if we replace m by m', by Fermat's Little Theorem. Therefore we have

$$\frac{(a^m - 1)B_m}{m} \equiv \frac{(a^{m'} - 1)B_{m'}}{m'} \mod p.$$

Taking a to be a primitive root mod p, we have $a^m - 1 \equiv a^{m'} - 1 \not\equiv 0 \mod p$, so we may cancel and conclude the result.

For $e > 1$, the argument is almost identical except for an additional complication which arises from the terms where $p|j$.

Start as above using Voronoi with $n = p^{t+e}$ to write

$$m^{-1}(a^m - 1)B_m \equiv a^{m-1} \sum_{j=1}^{p^{t+e}-1} j^{m-1}[ja/p^{t+e}] \mod p^e.$$

Let us break up the sum

$$\sum_{j=1}^{p^{t+e}-1} j^{m-1}[ja/p^{t+e}] = \sum_{j=1, p\not| j}^{p^{t+e}-1} j^{m-1}[ja/p^{t+e}] + \sum_{i=1}^{p^{t+e-1}-1} (pi)^{m-1}[ia/p^{t+e-1}].$$

The second term can be dealt with by using the same Voronoi formula with e replaced by $(e-1)$ and subtracting off, leaving us with the formula

$$\frac{(1 - p^{m-1})(a^m - 1)}{m} B_m \equiv a^{m-1} \sum_{j=1, p\not| j}^{p^{t+e}-1} j^{m-1}[ja/p^{t+e}] \mod p^e.$$

Now again the right hand is unchanged mod p^e if one replaces m by m', and one finishes exactly as in the case $e = 1$. □

We now introduce the important notion of a *regular prime*. Say that a prime p is *regular* if it does not divide the numerator of any of $B_2, B_4, ..., B_{p-3}$. Say p is *irregular* if it is not regular. The Kummer congruences have the following nice consequence.

Proposition 6.12. *There are infinitely many irregular primes.*

Proof. Suppose there are only finitely many, say $p_1, ..., p_t$, and set $M = N(p_1 - 1)...(p_t - 1)$ where N is to be chosen. We know that $|B_{2m}/2m| \to \infty$ as $m \to \infty$ so in particular if N is taken large enough, $|B_M/M| > 1$ so has some prime p dividing the numerator. But Clausen-von Staudt implies that each p_i divides the denominator, so $p_i \neq p$ and the same argument shows $p - 1 \not| M$ so we can find $0 < m < p - 1$ such that $m \equiv M \mod (p - 1)$. But this shows p is also irregular since

$$\frac{B_m}{m} \equiv \frac{B_M}{M} \equiv 0 \mod p.$$

□

7. Dirichlet characters, generalised Bernoulli numbers and L-Series

Recall that a *Dirichlet character* is a group homomorphism

$$\chi : (\mathbb{Z}/n\mathbb{Z})^* \to \mathbb{C}^*.$$

One can (abusing notation) associate to this a map

$$\chi : \mathbb{Z} \ni a \mapsto \begin{cases} 0 & \text{if } (a, n) \neq 1 \\ \chi(a) & \text{if } (a, n) = 1 \end{cases} \in \mathbb{C}.$$

The *conductor* $f = f(\chi)$ of a Dirichlet character is the n such that χ is given by a homomorphism $(\mathbb{Z}/f\mathbb{Z})^* \to \mathbb{C}^*$ and there exists no smaller $n'|f$ through which it factors $(\mathbb{Z}/f\mathbb{Z})^* \twoheadrightarrow (\mathbb{Z}/n'\mathbb{Z})^* \to \mathbb{C}^*$.

We should remark that of course a Dirichlet character of conductor f has image lying in $\mathbb{Q}(\zeta_{\phi(f)})^* \subset \mathbb{C}^*$, so one should view them as purely algebraic objects which are often given as embedded into $\mathbb{C}$.

For χ a Dirichlet character of conductor f, we now define the *generalised Bernoulli numbers* by the generating function

$$\sum_{n=0}^{\infty} B_{n,\chi} \frac{t^n}{n!} = \sum_{a=1}^{f} \frac{\chi(a) t e^{at}}{e^{ft} - 1}.$$

These need no longer be rational but it's easy to see they always lie in $\mathbb{Q}(\zeta_{\phi(f)})$. Since they are related, we introduce the *Bernoulli polynomials*

$$\sum_{n=0}^{\infty} B_n(X) \frac{t^n}{n!} = \frac{t e^{Xt}}{e^t - 1}.$$

Clearly $B_n(0) = B_n$, and more generally we see the following.

Lemma 7.1. *For all $n \geq 0$ we have the identity*

$$B_n(X) = \sum_{i=0}^{n} \binom{n}{i} B_i X^{n-i}.$$

Proof. This follows formally from comparing generating functions. Indeed

$$\begin{aligned}\sum_{n=0}^{\infty} B_n(X)\frac{t^n}{n!} &= \frac{te^{Xt}}{e^t-1} = e^{Xt}\frac{t}{e^t-1} \\ &= \sum_{k=0}^{\infty}\frac{X^k}{k!}t^k\frac{t}{e^t-1} \\ &= \sum_{k=0,l=0}^{\infty} X^k t^{k+l}\frac{B_l}{k!l!}.\end{aligned}$$

Comparing the t^n coefficients and multiplying through by $n!$ one gets

$$B_n(X) = \sum_{i=0}^{n} X^i B^{n-i}\binom{n}{i}$$

as required. □

In particular we observe that $B_n(X)$ is a polynomial with rational coefficients.

Proposition 7.2. *Let χ be a Dirichlet character and suppose N is any number divisible by the conductor f. Then we have the relation*

$$B_{n,\chi} = N^{n-1}\sum_{a=1}^{N}\chi(a)B_n(\frac{a}{N}).$$

Proof. Again, this comes down to a generating function computation

$$\begin{aligned}\sum_{n=0}^{\infty}\sum_{a=1}^{N}\chi(a)N^{n-1}B_n(a/N)\frac{t^n}{n!} &= \sum_{a=1}^{N}\sum_{n=0}^{\infty}\chi(a)\frac{1}{N}B_n(a/N)\frac{(Nt)^n}{n!} \\ &= \sum_{a=1}^{N}\chi(a)\frac{te^{(a/N)Nt}}{e^{Nt}-1} \\ &= \sum_{b=1}^{f}\sum_{c=0}^{N/f-1}\chi(b)\frac{te^{((b+cf)/N)Nt}}{e^{Nt}-1} \\ &= \sum_{b=1}^{f}\chi(b)te^{bt}\frac{\sum_{c=0}^{N/f-1}e^{cft}}{e^{Nt}-1} \\ &= \sum_{b=1}^{f}\chi(b)\frac{te^{bt}}{e^{ft}-1} \\ &= \sum_{n=0}^{\infty}B_{n,\chi}\frac{t^n}{n!}.\end{aligned}$$

□

For us the most important such numbers will be $B_{1,\chi}$, where the above formula can be given more explicitly as (for $\chi \neq 1$)

$$B_{1,\chi} = \frac{1}{f}\sum_{a=1}^{f} \chi(a)a.$$

With this expression, we are able to check the following very important congruence. Fix $\omega : (\mathbb{Z}/p\mathbb{Z})^* \to \mathbb{C}^*$ the *Teichmuller character* identifying $(\mathbb{Z}/p\mathbb{Z})^* \xrightarrow{\cong} \mu_{p-1} \subset \mathbb{C}^*$ in such a way that viewing $\mu_{p-1} \subset \mathbb{Q}(\zeta_{p-1})$ we have

$$\omega(a) \equiv a \mod p.$$

Theorem 7.3 (Second Kummer Congruences)**.** *Let $m \geq 2$ be even, and p prime such that $m \leq p-3$. Then (both sides are p-integral and)*

$$B_{1,\omega^{m-1}} \equiv \frac{B_m}{m} \mod p.$$

Proof. Work in $K = \mathbb{Q}(\zeta_{p-1})$, and let $\mathfrak{p}$ be a prime lying above p. Since $X^{p-1} - 1$ splits completely in $\mathbb{F}_p$, we have that $k(\mathfrak{p}) = \mathbb{F}_p$, and $(\mathcal{O}_K/\mathfrak{p}^2) = \mathbb{Z}/p^2\mathbb{Z}$. We claim that

$$\omega(a) \equiv a^p \mod \mathfrak{p}^2.$$

Indeed, this follows because for any $a \in \mathbb{Z}$, $(a^p)^{p-1} = a^{p(p-1)} \equiv 1 \mod p^2$, and also $a^p \equiv a \mod p$. Since these are the two properties that uniquely characterise $\omega(a)$, the claim follows. Since p doesn't ramify in K and $\mathfrak{p}$ was arbitrary, the Chinese remainder theorem implies that

$$\omega(a) \equiv a^p \mod p^2.$$

Equipped with this, we may write

$$pB_{1,\omega^{m-1}} = \sum_{a=1}^{p} \omega^{m-1}(a)a \equiv \sum_{a=1}^{p-1} a^{(m-1)p+1} = S_{(m-1)p+1}(p) \mod p^2.$$

But recall that (since $p-1 \not| m$)

$$S_{(m-1)p+1}(p) \equiv pB_{(m-1)p+1} \mod p^2.$$

Now finally the result follows from the first Kummer congruence: since $(m-1)p+1 \equiv m \mod (p-1)$ we have

$$B_{(m-1)p+1} \equiv ((m-1)p+1)^{-1}B_{(m-1)p+1} \equiv \frac{B_m}{m} \mod p.$$

□

Now, just as usual Bernoulli numbers are related to the Riemann zeta function, the generalised Bernoulli numbers are related to analytic objects called *Dirichlet L-functions*, which we now introduce.

For $Re(s) > 1$, let[8]

$$L(s, \chi) := \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} = \prod_p \frac{1}{1 - \chi(p)p^{-s}}.$$

For two examples, take $\chi = 1$ the trivial Dirichlet character, and of course $L(s, 1) = \zeta(s)$. One interesting feature of zeta is that as $s \to 1$, $\zeta(s) \to +\infty$ (because the harmonic series famously diverges).

On the other hand take $\chi : (\mathbb{Z}/3)^* \xrightarrow{\cong} \{\pm 1\} \subset \mathbb{C}^*$. This has L-series

$$L(s, \chi) = 1 - \frac{1}{2^s} + \frac{1}{4^s} - \frac{1}{5^s} + ...$$

By the alternating series test, this converges at $s = 1$ and in fact the function $L(s, \chi)$ will approach this limit as $s \to 1$. But there is certainly no hope once one takes $Re(s) < 1$. Our next task will be to show that one can make sense of (one can "analytically continue") the functions $L(s, \chi)$ for any $s \in \mathbb{C}$, by writing them as clever integrals instead.

We introduce the *Gamma function* which is defined by

$$\Gamma(s) = \int_0^{\infty} e^{-t} t^{s-1} dt.$$

This converges and defines an analytic function on $Re(s) > 0$.

Lemma 7.4 (Functional equation of the Gamma function)**.** *For $Re(s) > 1$, we have*

$$\Gamma(s) = (s - 1)\Gamma(s - 1).$$

Proof. This is an exercise in integration by parts. □

Equipped with the functional equation, it's easy to extend Γ to the entire complex plane. Indeed, let

$$\Gamma_k(s) = \frac{1}{s(s + 1)...(s + k - 1)} \Gamma(s + k)$$

and with the exception of simple poles at $s = 0, -1, ..., -(k - 1)$ it's clear that $\Gamma_k(s)$ is analytic on $Re(s) > -k$, and the functional equation tells us that if $Re(s) > -k_1 > -k_2$ then $\Gamma_{k_1}(s) = \Gamma_{k_2}(s)$.

We use this to extend the range of definition of $L(s, \chi)$ and prove the following theorem.

Theorem 7.5. *Let χ be a Dirichlet character. Then $L(s, \chi)$ admits an analytic continuation to the entire complex plane (except for a simple pole at $s = 1$ when χ is trivial), and for $m \geq 2$ an integer,*

$$L(1 - m, \chi) = -B_{m,\chi}/m.$$

[8]We brush over the nontrivial check that the product expression and the sum expression both really do converge and converge to the same thing.

For $Re(s) > 1$,

$$\chi(n)n^{-s}\Gamma(s) = \int_0^\infty \chi(n)e^{-nt}t^{s-1}dt.$$

Summing over all n (and exchanging a sum and an integral which in this case is not difficult)

$$L(s,\chi)\Gamma(s) = \int_0^\infty F_\chi(e^{-t})t^{s-1}dt$$

where

$$F_\chi(x) = \frac{1}{1-x^f}\sum_{a=1}^{f}\chi(a)x^a.$$

We cannot quite integrate by parts at this point because there is a pole at $t = 0$, so need a trick. Set $L^*(s,\chi) = (1-2^{1-s})L(s,\chi)$ and

$$R_\chi(x) = F_\chi(x) - 2F_\chi(x^2).$$

Substituting t for $2t$ into our equation above gives

$$2^{1-s}L(s,\chi)\Gamma(s) = 2\int_0^\infty F_\chi(e^{-2t})t^{s-1}.$$

Subtracting, we see that

$$L^*(s,\chi)\Gamma(s) = \int_0^\infty R_\chi(e^{-t})t^{s-1}dt.$$

To define our analytic continuation it helps to set $R_{\chi,0}(t) = R_\chi(e^{-t})$ and then

$$R_{\chi,n}(t) = \frac{dR_{\chi,n-1}}{dt},$$

and to note by explicit computation that $R_{\chi,n}(t)$ decays very rapidly as $t \to \infty$ and is bounded as $t \to 0$.

This allows us to integrate by parts, and obtain

$$\Gamma(s+k)L^*(s,\chi) = (-1)^k\int_0^\infty R_{\chi,k}(t)t^{s+k-1}dt.$$

This formula allows us to define $L*(s,\chi)$ on the whole complex plane, giving the required analytic continuation. We can also use it to compute special values. Recall we wish to compute $L(1-m,\chi)$ for $m \geq 2$ an integer, and substitute in $k = m, s = 1-m$, noting that $\Gamma(1) = 1$ by direct computation to get

$$(1-2^m)L(1-m,\chi) = (-1)^m\int_0^\infty R_{\chi,m}(t)dt = (-1)^{m-1}R_{\chi,m-1}(0),$$

where the second equality follows from the fundamental theorem of calculus.

But we can evaluate the right hand side via

$$\begin{aligned} R_\chi(e^{-t}) &= F_\chi(e^{-t}) - 2F_\chi(e^{-2t}) \\ &= \frac{1}{1-e^{-tf}}\sum_{a=1}^{f}\chi(a)e^{-ta} - \frac{2}{1-e^{-2tf}}\sum_{a=1}^{f}\chi(a)e^{-2ta} \\ &= \frac{1}{t}\sum_{k=1}^{\infty}(-1)^k(1-2^k)(B_{k,\chi}/k!)t^k. \end{aligned}$$

From this we conclude that

$$L(1-m,\chi) = -B_{m,\chi}/m.$$

8. The Analytic Class Number Formula

In this section we make a brief excursion to prove a very important general fact about number fields relating their arithmetic invariants to special values of their zeta functions. We are loosely following the online notes of Gary Sivek.

Let K be a number field. If $\mathfrak{a} \subset \mathcal{O}_K$ is an ideal, we define its *norm* $N(\mathfrak{a}) = |\mathcal{O}_K/\mathfrak{a}|$.

Lemma 8.1. *If $\alpha \in \mathcal{O}_K$,*

$$N((\alpha)) = |N_{K/\mathbb{Q}}(\alpha)|.$$

Proof. Since $N_{K/\mathbb{Q}}(\alpha)$ is the determinant of α viewed as a $\mathbb{Q}$-linear automorphism of K, its absolute value is the volume of a parallelopiped in K after being multiplied by α divided by its original volume, which is exactly the index $(\mathcal{O}_K : (\alpha))$. □

The *Dedekind zeta function* $\zeta_K(s)$ of the field K is given by the formula

$$\zeta_K(s) = \sum_{\mathfrak{a}\subset\mathcal{O}_K}\frac{1}{N(\mathfrak{a})^s},$$

where $\mathfrak{a}$ runs over all nonzero ideals of $\mathcal{O}_K$.

A key trick for studying $\zeta_K(s)$ will be breaking the sum up as

$$\zeta_K(s) = \sum_{C\in Cl(K)}\sum_{\mathfrak{a}\in C}\frac{1}{N(\mathfrak{a})^s}.$$

It will help to introduce the notation

$$f_C(s) = \sum_{\mathfrak{a}\in C}\frac{1}{N(\mathfrak{a})^s}$$

for these partial sums.

Now, consider any ideal $\mathfrak{b} \in C^{-1}$. Multiplication by $\mathfrak{b}$ gives a bijection

$$\{\text{ideals in } C\} \cong \{\text{principal ideals divisible by } \mathfrak{b}\}.$$

Using this we can compute using elements of $\mathcal{O}_K$

$$f_C(s) = N(\mathfrak{b})^s \sum_{(\alpha)\subset\mathfrak{b}} \frac{1}{|N(\alpha)|^s}.$$

To study these sums, it will be useful to first make the following abstract geometric analysis.

Proposition 8.2. *Let X be a cone*[9] *in $\mathbb{R}^n$, $F : X \to \mathbb{R}_{>0}$ a function such that for $\lambda > 0$, $f(\lambda x) = \lambda^n x$ and $B = \{x \in X | F(x) \leq 1\}$ is bounded with volume $v = Vol(B) > 0$. Suppose we also have a lattice Γ with covolume $\delta = Vol(\mathbb{R}^n/\Gamma) < \infty$. Then the sum*

$$\zeta_{F,\Gamma}(s) = \sum_{x\in\Gamma\cap X} \frac{1}{F(x)^s}$$

converges for $\Re(s) > 1$ and

$$\lim_{s\to 1}(s-1)\zeta_{F,\Gamma}(s) = \frac{v}{\delta}.$$

Proof. Let us enumerate $\Gamma \cap X = \{x_1, x_2, ...\}$ where $F(x_1) \leq F(x_2) \leq .....$ Our key estimate will be that for any $\epsilon > 0$ there is k_0 such that for all $k \geq k_0$ and $s \geq 1$

$$\left(\frac{v}{\delta} - \epsilon\right)^s \frac{1}{k^s} < \frac{1}{F(x_k)^s} < \left(\frac{v}{\delta} + \epsilon\right)^s \frac{1}{k^s}.$$

Let us establish this claim. Note that multiplication by r establishes a bijection

$$\gamma(r) := |\frac{1}{r}\Gamma \cap B| = |\{x \in \Gamma \cap X : F(x) \leq r^n.\}|.$$

On the one hand

$$\lim_{r\to\infty} \frac{\gamma(r)}{r^n} = \lim_{r\to\infty} \frac{|\frac{1}{r}\Gamma \cap B|}{r^n} = \frac{v}{\delta}.$$

Now let $r_k = F(x_k)^{1/n}$. It is clear that for all $\eta > 0$,

$$\gamma(r_k - \eta) < k \leq \gamma(r_k).$$

Let us re-write this as

$$\frac{\gamma(r_k-\eta)}{(r_k-\eta)^n}\frac{(r_k-\eta)^n}{r_k^n} < \frac{k}{F(x_k)} \leq \frac{\gamma(r_k)}{r_k^n}.$$

As $k \to \infty$, $r_k \to \infty$ and we see that

$$\lim_k \frac{k}{F(x_k)} = \frac{v}{\delta},$$

which implies the required estimate after raising to the s-th power and rearranging.

To see why the proposition now follows, write

$$\zeta_{F,\Gamma,k_0}(s) = \sum_{k\geq k_0} \frac{1}{F(x_k)^s},$$

[9]In this context a cone is any subset closed under $\mathbb{R}_{>0}$-multiplication.

and note that the estimate gives

$$\left(\frac{v}{\delta}-\epsilon\right)^s \sum_{k\geq k_0}\frac{1}{k^s} < \zeta_{F,\Gamma,k_0}(s) < \left(\frac{v}{\delta}+\epsilon\right)^s \sum_{k\geq k_0}\frac{1}{k^s}.$$

In particular, convergence for $\Re(s) > 1$ follows from that of the Riemann zeta function, and the claim about the residue at $s = 1$ from that it only depends on the tail, and that for the Riemann zeta function $\lim_{s\to 1}(s-1)\zeta(s) = 1$. □

Now we apply this setup to the study of the $f_C(s)$. The idea is to specify a cone such that for each (α) there is a unique generator α lying in the cone (i.e. such that multiplying by any nontrivial unit takes one outside the cone).

Let us first give a general setup. Suppose K has r_1 real places $\tau_1, ..., \tau_{r_1}$ and r_2 pairs $\tau_{r_1+1}, \bar{\tau}_{r_1+1}, ..., \tau_{r_1+r_2}, \bar{\tau}_{r_1+r_2}$ of complex places. We have $n = [K : \mathbb{Q}] = r_1 + 2r_2$ and have the canonical injective ring homomorphism $\tau : K \to \mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$. We also have a "log of absolute value" map $l : (\mathbb{R}^{r_1} \times \mathbb{C}^{r_2})^* \to \mathbb{R}^{r_1+r_2}$ given by

$$(x_1, ..., x_{r_1}, z_{r_1+1}, ..., z_{r_1+r_2}) \mapsto (\log|x_1|, ..., \log|x_{r_1}|, 2\log|z_{r_1+1}|, ..., 2\log|z_{r_1+r_2}|)$$

and only defined on the multiplicative subgroup of elements with no zero components.

Particularly significant is the composite $\phi = l \circ \tau : K^* \to \mathbb{R}^{r_1+r_2}$

$$\alpha \mapsto (\log|\tau_1(\alpha)|, ..., \log|\tau_{r_1}(\alpha)|, 2\log|\tau_{r_1+1}(\alpha)|, ..., 2\log|\tau_{r_1+r_2}(\alpha)|).$$

Recall Dirichlet's unit theorem which says that the unit group $\mathcal{O}_K^*$ has rank $r_1 + r_2 - 1$, which (together with the fact the kernel of ϕ consists of roots of unity) implies that $\phi(\mathcal{O}_K^*)$ is a lattice in the subspace $\{x_1 + ... + x_{r_1+r_2} = 0\} \subset \mathbb{R}^{r_1+r_2}$. Let $\epsilon_1, ..., \epsilon_{r_1+r_2-1} \in \mathcal{O}_K^*$ be a basis for the free part of $\mathcal{O}_K^*$ (which we recall is often called a basis of *fundamental units*). We also will let $\omega_K = |Tors(\mathcal{O}_K^*)|$ be the order of the cyclic group of roots of unity in $\mathcal{O}_K^*$.

We also let $\lambda = (1, 1, ..., 1, 2, 2, ..., 2) \in \mathbb{R}^{r_1+r_2}$ be the image of a hypothetical e under ϕ. Then

$$\{\lambda, \phi(\epsilon_1), ..., \phi(\epsilon_{r_1+r_2-1})\}$$

is a convenient basis for $\mathbb{R}^{r_1+r_2}$. For any $x \in (\mathbb{R}^{r_1} \times \mathbb{C}^{r_2})^*$ we may write

$$l(x) = c\lambda + c_1\phi(\epsilon_1) + ... + c_{r_1+r_2-1}\phi(\epsilon_{r_1+r_2-1}).$$

Note that always $c = \log|N(x)|/n$.

Now let X be the cone in $\mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$ given by those x defined by the constraints

- We insist $x \in (\mathbb{R}^{r_1} \times \mathbb{C}^{r_2})^*$, giving it a meaningful logarithm.
- We pin down the free part of the units by insisting that for all $1 \leq i \leq n-1$,

$$0 \leq c_i < 1.$$

- We pin down the root of unity by

$$0 \leq arg(x_1) < \frac{2\pi}{\omega_K}.$$

This is a cone. Indeed if $r > 0$ and $x \in X$, of course $rx \in (\mathbb{R}^{r_1} \times \mathbb{C}^{r_2})^*$,

$$l(rx) = (\log r)\lambda + l(x),$$

and $arg(rx_1) = arg(x_1)$.

Lemma 8.3. *Let $\alpha \in \mathcal{O}_K$ be nonzero. There is a unique unit u such that $\tau(u\alpha) \in X$.*

Proof. This is true by design. Since $\alpha \neq 0$, $\tau(\alpha) \in (\mathbb{R}^{r_1} \times \mathbb{C}^{r_2})^*$. Now look at $\phi(\alpha)$. There are clearly unique $m_1, ..., m_{r_1+r_2-1} \in \mathbb{Z}$ such that replacing α by $\epsilon_1^{m_1}...\epsilon_{r_1+r_2-1}^{m_{r_1+r_2-1}}\alpha$ we have $0 \leq c_i < 1$, and there's a unique root of unity by which we can multiply this to force $0 \leq arg(\alpha) < \frac{2\pi}{\omega_K}$. □

Using this lemma, we may re-write

$$f_C(s) = N(\mathfrak{b})^s \sum_{\alpha \in \tau(\mathfrak{b}) \cap X} \frac{1}{|N(\alpha)|^s}.$$

Written this way, it is computable using the main proposition. Indeed we can already say the following.

Theorem 8.4. *The Dedekind zeta function $\zeta_K(s)$ given as a Dirichlet series converges absolutely for $\Re(s) > 1$.*

Proof. Since $\mathcal{O}_K$ is a lattice in $K \otimes_{\mathbb{Q}} \mathbb{R}$, and $\mathfrak{b}$ has finite index in $\mathcal{O}_K$ it is also a lattice and so $f_C(s)$ converges absolutely for $\Re(s) > 1$ by the proposition (taking $F(x) = |N(x)|$ in the obvious sense). Since the class group is finite, $\zeta_K(s)$ is a finite sum of such functions and so also converges absolutely. □

We would like to also use the proposition to compute the residue at $s = 1$. To do this, we must compute v and δ for the cone X and the lattice $\tau(\mathfrak{b})$.

Lemma 8.5. *The lattice $\tau(\mathfrak{b}) \subset \mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$ has covolume*

$$\delta = N(\mathfrak{b})|\delta_K|^{1/2}.$$

Proof. Let $x_1, ..., x_n \in \mathfrak{b}$ be a $\mathbb{Z}$-basis for $\mathfrak{b}$. Then

$$Covol(\tau(\mathfrak{b}))^2 = |\det \tau_j(x_i)^2| = |\Delta(x_1, ..., x_n)| = N(\mathfrak{b})^2|\delta_K|.$$

□

Before we can state the formula for the volume v we will need to define a constant called the *regulator* which measures the size of the fundamental units (since X is defined using them, that such a constant appears should not be a surprise). There are $r_1 + r_2$ places and $r_1 + r_2 - 1$ fundamental units, but any unit has norm 1, so if we forget one of the places we can always recover its value from the others. We therefore define the regulator to be (forgetting the place $\tau_{r_1+r_2}$) the determinant of the $r_1 + r_2 - 1$-dimensional square matrix

$$R_K = |\det(\log|\tau_j(\epsilon_i)|)|.$$

By easy arguments using row operations one sees that this doesn't depend on which place we dropped or on the choice of basis ϵ_i.

Lemma 8.6. *The ball* $B = \{x \in X || N(x)| \leq 1\}$ *has volume*

$$v = 2^{r_1+r_2}\pi^{r_2}\frac{R_K}{\omega_K}.$$

Proof. First we drop the constraint on argument replacing B by $B_1 = B \cup \zeta B \cup ... \cup \zeta^{\omega_K - 1}B$ which is a disjoint union of regions with the same volume multiplying the number to compute by ω_K, and then focus our attention on $B' = \{b \in B_1 | b_i > 0 \ \forall i \leq r_1\}$, which cuts the volume down by 2^{r_1}. This reduces us to showing

$$Vol(B') = (2\pi)^{r_2}R_K.$$

Now change co-ordinates from $(b_i) \in \mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$ to $(\rho_i, \theta_i) \in \mathbb{R}^{r_1} \times (\mathbb{R}^{r_2} \times \mathbb{R}^{r_2})$. where $\rho_i = |b_i|$ and $\theta_i = arg(b_i)$ (where we only have θ_i when i indexes a complex place). Now B is cut out by the inequalities

$$0 < \rho_i, \quad \prod_i \rho_i^{\lambda_i} \leq 1, \quad 0 \leq c_i < 1.$$

Recalling that

$$l(b) = \frac{|N(b)|}{n}\lambda + \sum_i c_i \phi(\epsilon_i),$$

we can look at the j-th component $l(b)_j$ and get the equation

$$\lambda_j \log \rho_j = \frac{\lambda_j}{n} \log \prod \rho_k^{\lambda_k} + \sum_i c_i \phi_j(\epsilon_i).$$

Hence we can recover ρ_j from c_i and $c = \prod \rho_k^{\lambda_k}$, so again let's make the change of variables from (ρ_i, θ_i) to (c, c_i, θ_i). Now the constraints just are $0 < c \leq 1, 0 \leq c_i < 1, 0 \leq \theta_i < 2\pi$ so we have a box of volume $(2\pi)^{r_2}$. The computation therefore reduces to showing the total Jacobian between (b_i) and this new co-ordinate system is R_K.

It is straighforward that

$$dc_1....dc_{r_1}|dc_{r_1+1}|^2...|dc_{r_1+r_2}|^2 = 2^{r_2}\rho_{r_1+1}...\rho_{r_1+r_2}d\rho_1....d\rho_{r_1+r_2}d\theta_{r_1+1}...d\theta_{r_1+r_2}.$$

The more interesting computation is that of the ratio $J = d\rho_1...d\rho_{r_1+r_2}/dcdc_1...dc_{r_1+r_2-1}$. We have

$$\partial\rho_i/\partial c = \frac{\rho_j}{nc}$$

and

$$\partial\rho_i/\partial c_j = \frac{\rho_i}{\lambda_i}\phi_i(\epsilon_j).$$

Therefore by a short matrix computation

$$J = \prod_i \rho_i . \frac{1}{nc2^{r_2}} nR_K = \frac{R_K}{2^{r_2}\prod_{i \geq r_1+1}\rho_i}.$$

We conclude that

$$\begin{aligned} Vol(B') &= \int_{B'} dc_1...dc_{r_1}|dc_{r_1+1}|^2...|dc_{r_1+r_2}|^2 \\ &= 2^{r_2}\int_{B'} \rho_{r_1+1}...\rho_{r_1+r_2}d\rho_1....d\rho_{r_1+r_2}d\theta_{r_1+1}...d\theta_{r_1+r_2} \\ &= 2^{r_2}\int_{B'} \rho_{r_1+1}...\rho_{r_1+r_2}\frac{R_K}{2^{r_2}\prod_{i\geq r_1+1}\rho_i}dcdc_1...dc_{r_1+r_2-1}d\theta_{r_1+1}...d\theta_{r_1+r_2} \\ &= (2\pi)^{r_2}R_K. \end{aligned}$$

□

Putting all the pieces together, we conclude the following amazing theorem.

Theorem 8.7 (Class Number Formula)**.** *The limit of* $(s-1)\zeta_K(s)$ *as* $s\to 1$ *exists and is given by*

$$\lim_{s\to 1}(s-1)\zeta_K(s) = h_K\frac{2^{r_1}(2\pi)^{r_2}R_K}{\omega_K|d_K|^{1/2}}.$$

This formula tells us (roughly) that if we know two of:

- The residue of $\zeta_K(s)$ at $s=1$,
- The regulator R_K (in practice, perhaps a basis of fundamental units),
- The order h_K of the class group of K,

then we are able to determine the third. Our approach to Kummer's theorem will involve a trick to eliminate having to compute the regulator, and then computing the necessary residue in terms of Bernoulli numbers.

9. Applying class number formulas to cyclotomic fields

We now specialise to the case where $K\subset\mathbb{Q}(\zeta_n)$ for some n. By Galois theory, such K corresponds to a quotient $(\mathbb{Z}/n\mathbb{Z})^*\twoheadrightarrow\Gamma = Gal(K/\mathbb{Q})$. Here the Dedekind zeta function can be related directly to Dirichlet L-functions.

We begin by making explicit the filtration $1\subset I_p\subset D_p\subset\Gamma$ associated with a prime p.

Lemma 9.1. (1) *For* $K=\mathbb{Q}(\zeta_n)$, $\Gamma\cong(\mathbb{Z}/n\mathbb{Z})^*\cong(\mathbb{Z}/p^k\mathbb{Z})^*\times(\mathbb{Z}/n'\mathbb{Z})^*$ *with* $n=p^kn', p\not| n'$, *we have*

$$I_p = (\mathbb{Z}/p^k\mathbb{Z})^*,\quad D_p = (\mathbb{Z}/p^k\mathbb{Z})^*\times\langle[p]\rangle.$$

(2) *For* $K\subset\mathbb{Q}(\zeta_n)$, *corresponding to* $\pi:(\mathbb{Z}/n\mathbb{Z})^*\to\Gamma$, I_p *and* D_p *are the images of* $(\mathbb{Z}/p^k\mathbb{Z})^*$ *and* $(\mathbb{Z}/p^k\mathbb{Z})^*\times\langle[p]\rangle$ *under* π.

Proof. For (1), recall that p does not ramify in $\mathbb{Q}(\zeta_{n'})/\mathbb{Q}$, but the ramification index of p in $\mathbb{Q}(\zeta_n)$ is at least that of p in $\mathbb{Q}(\zeta_{p^k})$ which is $p^{k-1}(p-1)=|(\mathbb{Z}/p^k\mathbb{Z})^*|$. It follows immediately that

$$I_p = Ker((\mathbb{Z}/n\mathbb{Z})^*\to(\mathbb{Z}/n'\mathbb{Z})^*) = (\mathbb{Z}/p^k\mathbb{Z})^*.$$

For D_p, note first that $[p]$ corresponds to the automorphism $\zeta \mapsto \zeta^p$ which acting as Frobenius in characteristic p acts on each factor of $\mathcal{O}_K/p\mathcal{O}_K = \prod \mathcal{O}_K/\mathfrak{p}_i^e$ individually, so in particular $[p] \in D_p$. However, $[p]$ also induces the Frobenius automorphism on a residue field $\mathcal{O}_K/\mathfrak{p}_i$ so in particular its image in $Gal((\mathcal{O}_K/\mathfrak{p}_i)/\mathbb{F}_p)$ generates the Galois group, which has order f. Thus we see that $(\mathbb{Z}/p^k\mathbb{Z})^* \times \langle [p] \rangle \subset D_p$ has order at least ef, which implies the equality required.

Now let us prove (2). Firstly, if $\sigma \in (\mathbb{Z}/n\mathbb{Z})^*$ fixes a prime $\mathfrak{p}$ of $\mathbb{Q}(\zeta_n)$ its image in Γ certainly fixes the unique prime that $\mathfrak{p}$ divides, and similarly for acting trivially on the residue field, so we have

$$\pi((\mathbb{Z}/p^k\mathbb{Z})^* \times \langle [p] \rangle) \subset D_p, \quad \pi((\mathbb{Z}/p^k\mathbb{Z})^*) \subset I_p.$$

The second inclusion is an equality because viewing $\mathbb{Q}(\zeta_n)$ as the compositum of $\mathbb{Q}(\zeta_{p^k})$ and $\mathbb{Q}(\zeta_{n'})$ we can see that $|\pi((\mathbb{Z}/p^k\mathbb{Z})^*)| = [K \cap \mathbb{Q}(\zeta p^k) : \mathbb{Q}] = e_{p,K}$. With this in hand, the first inclusion is an equality because the image of $[p]$ is still the Frobenius at p for $K \cap \mathbb{Q}(\zeta_{n'})/\mathbb{Q}$. □

Proposition 9.2. *Let $K \subset \mathbb{Q}(\zeta_n)$ be a subfield of a cyclotomic field as above, and $\Gamma = Gal(K/\mathbb{Q})$.*

(1) *The set X of Dirichlet characters $(\mathbb{Z}/n\mathbb{Z})^* \to \mathbb{C}^*$ which factor through Γ form a group and the natural pairing*

$$\Gamma \times X \to \mathbb{C}^*$$

is a perfect pairing.

(2) *Let p be any prime. The sets $Y = \{\chi \in X | \chi(p) \neq 0\}$ and $Z = \{\chi \in X | \chi(p) = 1\}$ are both subgroups of X and if $p\mathcal{O}_K = (\mathfrak{p}_1 \ldots \mathfrak{p}_r)^e$ with common degree f we have*

$$e = [X : Y], f = [Y : Z], r = |Z|.$$

Proof. Given $\chi_1, \chi_2 \in X$ it's clear that the product $\chi_1\chi_2^{-1}$ is also a Dirichlet character and still factors through Γ so lies in X, proving the first claim. For the second it suffices to check that as a group $X = Hom_{Ab-Gp}(\Gamma, \mathbb{C}^*)$, which is obvious.

For (2), since Γ is abelian there is a well-defined decomposition group D_p and inertia group I_p, giving a filtration

$$1 \subset I_p \subset D_p \subset \Gamma.$$

Writing $H^\perp = \{x \in X | \forall h \in H \;\; x(h) = 1\}$ for a subgroup $H \subset \Gamma$, one sees that the perfect duality between Γ and X induces a perfect duality between $H^\perp$ and Γ/H. Thus (2) will follow if we can check that $Y = I_p^\perp$ and $Z = D_p^\perp$.

But this is easy given the previous lemma. Indeed, the condition $\chi(p) \neq 0$ exactly says that χ factors through $(\mathbb{Z}/n\mathbb{Z})^* \to (\mathbb{Z}/n'\mathbb{Z})^*$, so it kills $I_p = \pi((\mathbb{Z}/p^k\mathbb{Z})^*)$. Thus $Y = I_p^\perp$. Also by the above lemma, the additional condition that $\chi(p) = 1$ is precisely what is needed to kill D_p. □

Proposition 9.3. *We have the formula*

$$\zeta_K(s) = \prod_{\chi : \Gamma \to \mathbb{C}^*} L(s, \chi)$$

where the product is over all Dirichlet characters χ of conductor dividing n which factor through Γ (where we count each only once: for example the trivial character viewed after projection from $(\mathbb{Z}/n\mathbb{Z})^*$ doesn't contribute an additional factor).

Proof. Both sides can be written as infinite products of Euler factors, so it suffices to show

$$\prod_{\mathcal{P}|p}(1-N(\mathcal{P})^{-s}) = \prod_{\chi:\Gamma\to\mathbb{C}^*}(1-\chi(p)p^{-s}).$$

Since $K/\mathbb{Q}$ is Galois, $p\mathcal{O}_K = (\mathfrak{p}_1...\mathfrak{p}_r)^e$ with each $\mathfrak{p}_i$ of some fixed degree f. Therefore

$$\prod_{\mathcal{P}|p}(1-N(\mathcal{P})^{-s}) = (1-p^{-sf})^r.$$

On the other hand, let X be the group of Dirichlet characters factoring through Γ. By the previous proposition, it's clear that (noting we may discard all χ with $\chi(p)=0$)

$$\begin{aligned}\prod_{\chi\in X}(1-\chi(p)p^{-s}) &= \prod_{\chi\in Y/Z}(1-\chi(p)p^{-s})^r \\ &= \prod_{a=1}^{f}(1-\zeta_f^a p^{-s})^r \\ &= (1-p^{-sf})^r.\end{aligned}$$

Comparing the two equalities, we get the result. □

Since we wish to use the class number formula, the key question is that of the behaviour of these L-functions at $s=1$.

Lemma 9.4. *The L-function $L(s,\chi)$ is defined on $\mathbb{C}$. It has no pole at $s=1$.*

Proof. One way to do this would be to check by evaluating an integral that $L^*(1,\chi)=0$. We give a different method which also gives a different way to analytically continue $L(s,\chi)$ to $Re(s)>0$.

Write

$$L(s,\chi)=\sum_{n=1}^{\infty}\chi(n)n^{-s}=\sum_{n=0}^{\infty}(\sum_{m=1}^{n}\chi(m))(n^{-s}-(n+1)^{-s}).$$

Since the sums over Dirichlet values are bounded (say by the conductor f) and by the binomial theorem

$$n^{-s}-(n+1)^{-s} = sn^{-(s+1)} + \text{higher order terms}$$

this expression converges for all $Re(s)>0$. In particular it converges at $s=1$. □

The relationship between L-series and Dedekind zeta functions gives the following important result.

Proposition 9.5. *Let $\chi\neq 1$ be a nontrivial Dirichlet character. Then*

$$L(1,\chi)\neq 0.$$

Proof. Let $K = \mathbb{Q}(\zeta_{f_\chi})$. By the class number formula, in particular we know that $\zeta_K(s)$ has a simple pole at $s = 1$. But also by the previous proposition

$$\zeta_K(s) = \prod_{\chi:(\mathbb{Z}/f\mathbb{Z})^* \to \mathbb{C}^*} L(s, \chi)$$

and we know each $L(s, \chi)$ converges at $s = 1$ except $L(s, 1) = \zeta(s)$ which also has a simple pole. Since $ord_{s=1}\zeta = ord_{s=1}\zeta_K = -1$, and $ord_{s=1}L(s, \chi) \geq 0$ for $\chi \neq 1$, we conclude that no factor $L(1, \chi)$ can vanish. □

Having obtained this result so cleanly, it is worth noting its most famous application.

Theorem 9.6 (Dirichlet)**.** *Let $a, m \geq 1$ be coprime positive integers. There are infinitely many primes of the form*

$$p = km + a.$$

Proof. The point is one can pick out residue classes using the fact that

$$\sum_{\chi:(\mathbb{Z}/m\mathbb{Z})^* \to \mathbb{C}^*} \chi(a) = \begin{cases} \phi(m) & \text{if } a \equiv 1 \mod m \\ 0 & \text{otherwise.} \end{cases}$$

Let us combine this with the identity

$$\log L(s, \chi) = -\sum_{p\text{prime}} \log(1 - \chi(p)p^{-s}) = \sum_p \frac{\chi(p)}{p^s} + g_\chi(s)$$

where $g_\chi(s)$ is holomorphic for $Re(s) > 1/2$, to get

$$\sum_\chi \chi(a^{-1}) \log L(s, \chi) = \sum_\chi \sum_p \chi(a^{-1}p)/p^s + g(s) = \sum_{p \equiv a \bmod m} \phi(m)/p^s + g(s),$$

where $g(s)$ is also holomorphic for $Re(s) > 1/2$. In particular observe that if this sum diverges as $s \to 1$ there must be infinitely many terms on the right hand side.

But the left hand side does diverge: the terms corresponding to $\chi \neq 1$ are bounded because $L(1, \chi) \neq 0$, and the term corresponding to $\chi = 1$ tends to $+\infty$. □

We wish to apply class number formulas, so let's now turn to the question of evaluating $L(1, \chi)$. Recall that since $1 - 2^{1-s}$ is nonzero at $s = 0$, we were able to evaluate $L(0, \chi)$ using integration by parts, and in fact

$$L(0, \chi) = -B_{1,\chi}.$$

We will make use of the following identity, which is the *functional equation* of $L(s, \chi)$. We omit the proof, which can be found in any standard text on analytic number theory.

Theorem 9.7 (Functional equation for Dirichlet L-functions)**.** *Let χ be a Dirichlet character of conductor f. Let $r \in \{0, 1\}$ be such that $\chi(-1) = (-1)^r$, and $\tau(\chi)$ be the Gauss sum*

$$\tau(\chi) = \sum_{a=1}^{f} \chi(a) e^{2\pi i(a/f)}.$$

Then we have the identity

$$\left(\frac{\pi}{f}\right)^{-(1-s+r)/2}\Gamma\left(\frac{1-s+r}{2}\right)L(1-s,\bar{\chi})=\frac{i^r f^{1/2}}{\tau(\chi)}\left(\frac{\pi}{f}\right)^{-(s+r)/2}\Gamma\left(\frac{s+r}{2}\right)L(s,\chi).$$

If $r=1$, plugging in $s=1$ works out very nicely and one gets the following.

Corollary 9.8. *Suppose $\chi(-1)=-1$. Then*

$$L(1,\chi)=\tau(\chi)\frac{\pi}{if}L(0,\bar{\chi})=\tau(\chi)\frac{i\pi}{f}B_{1,\bar{\chi}}.$$

Note that to prove this one needs the computation

$$\Gamma(1/2)=\int_0^\infty x^{1/2}e^{-x}dx=\int_0^\infty 2e^{-u^2}du=\sqrt{\pi}.$$

Suppose $K\subset\mathbb{Q}(\zeta_n)$ is a field and X is the group of Dirichlet characters which factor through $Gal(K/\mathbb{Q})$. Since $\lim_{s\to 1}(s-1)\zeta(s)=1$, the class number formula combined with our theorem on the factorisation of the Dedekind zeta function gives

$$\frac{2^{r_1}(2\pi)^{r_2}h_K R_K}{\omega_K\sqrt{|\delta_K|}}=\prod_{\chi\in X,\chi\neq 1}L(1,\chi).$$

In particular, letting $K=\mathbb{Q}(\zeta_n)$ and $K^+=\mathbb{Q}(\zeta_n+\zeta_n^{-1})$ (and we use the helpful abbreviations $h=h_K, h^+=h_{K^+}$), and dividing through by a power of π we have the following.

Proposition 9.9 (Relative class number formula: preliminary form)**.** *We have the equality of complex numbers*

$$\frac{h}{h^+}\frac{2R_K}{\omega_K R_{K^+}}\sqrt{\frac{|\delta_{K^+}|}{|\delta_K|}}=\prod_{\chi:(\mathbb{Z}/n\mathbb{Z})^*\to\mathbb{C}^*|\chi(-1)=-1}\tau(\chi)\frac{i}{f}B_{1,\chi}.$$

This formula should be striking: it is our first equation directly relating Bernoulli numbers to class numbers. We now aim to simplify some of the factors involved.

Firstly, the factors involving discriminants, conductors, i, and Gauss sums are arguably the simplest, and can be simplified using the following proposition.

Proposition 9.10 (Conductor-Discriminant formula)**.** *Let $K\subset\mathbb{Q}(\zeta_n)$ be a subfield of a cyclotomic field (and as always K has r_2 pairs of complex places), X the group of Dirichlet characters factoring through $Gal(K/\mathbb{Q})$. Then*

$$\delta_K=(-1)^{r_2}\prod_{\chi\in X}f_\chi$$

and

$$\prod_{\chi\in X}\tau(\chi)=i^{r_2}\sqrt{|\delta_K|}.$$

Proof. This follows by comparing the functional equation for Dirichlet L-series with that for the Dedekind zeta function, and the formula

$$\zeta_K(s) = \prod_{\chi \in X} L(s, \chi).$$

Indeed, setting $A = 2^{-r_2}\pi^{-[K:\mathbb{Q}]/2}\sqrt{|\delta_K|}$, the functional equation for $\zeta_K(s)$ is

$$A^s\Gamma(s/2)^{r_1}\Gamma(s)^{r_2}\zeta_K(s) = A^{1-s}\Gamma((1-s)/2)^{r_1}\Gamma(1-s)^{r_2}\zeta_K(1-s).$$

Since $K/\mathbb{Q}$ is Galois there are only two cases.

If K is real, $r_1 = [K:\mathbb{Q}], r_2 = 0$, and $\chi(-1) = 1$ for all $\chi \in X$ (since χ factors through $Gal(\mathbb{Q}(\zeta_n + \zeta_n^{-1})$ which is exactly the group generated by -1). Comparing functional equations and squaring (for convenience) gives the following two equations.

Looking at the factor being raised to power s,

$$A^2 = \prod_{\chi} \frac{f_\chi}{\pi}$$

and the constant term gives

$$1 = \prod_{\chi} \frac{f_\chi^{1/2}}{\tau(\chi)}.$$

We conclude that

$$|\delta_K| = \prod_{\chi \in X} f_\chi = \prod_{\chi} (\tau(\chi))^2,$$

from which both parts of the proposition follow.

If K is complex, the argument is similar, but since half the L-factors now come from odd characters we get complications in the gamma factors arising, and need the analytic identity of Whittaker and Watson

$$\Gamma(s/2)\Gamma((s+1)/2) = 2^{1-s}\sqrt{\pi}\Gamma(s).$$

With this in hand, we can simplify the "s" side of the functional equation

$$A^s\Gamma(s)^{r_2} = \prod_{\chi \text{ even}} \Gamma(s/2)\frac{f_\chi^{1/2}}{\tau(\chi)}(f_\chi/\pi)^{s/2} \prod_{\chi \text{ odd}} \Gamma((s+1)/2)\frac{if_\chi^{1/2}}{\tau(\chi)}(f_\chi/\pi)^{s/2}$$

to

$$A^s = i^{r_2} \prod_{\chi} \frac{f_\chi^{1/2}}{\tau(\chi)}(f_\chi/2\pi)^{s/2}.$$

Again, the $(-)^s$-part of this gives

$$|\delta_K| = \prod_{\chi} f_\chi.$$

This together with the usual observation about signs of discriminants proves the first part of the proposition.

For the second, again look at the constant term and we see that

$$\prod_\chi \tau(\chi) = i^{r_2} \prod_\chi f_\chi^{1/2} = i^{r_2}\sqrt{|\delta_K|}.$$

□

The consequence of this is that in the relative class number formula we see that

$$\sqrt{\frac{|\delta_K|}{|\delta_{K^+}|}} = i^{-\phi(n)/2} \prod_{\chi \text{ odd}} \tau(\chi) = i^{-\phi(n)/2} \prod_{\chi \text{ odd}} f_\chi^{1/2}.$$

We may therefore cancel many factors from the relative class number formula, reducing it to

$$\frac{h}{h^+}\frac{2R_K}{\omega_K R_{K^+}} = (-1)^{r_2} \prod_{\chi \text{ odd}} B_{1,\chi}.$$

The next factor we turn to will be R_K/R_{K^+}. The first thing to observe is that since $\mathcal{O}_{K^+}^* \subset \mathcal{O}_K^*$ is a subgroup and both are abstractly finitely generated abelian groups, both with free rank $\phi(n)/2 - 1$, $\mathcal{O}_{K^+}^*$ actually sits as a finite index subgroup of $\mathcal{O}_K^*$. This alone tells us that the a priori highly transcendental number R_K/R_{K^+} is in fact a rational number. Even more miraculously, we are able to compute it.

Lemma 9.11. (1) *Let K be a number field, and $U \subset \mathcal{O}_K^*/\mu_K$ a finite index subgroup of the unit group modulo torsion. Take $\eta_1, ..., \eta_{r_1+r_2-1}$ a $\mathbb{Z}$-basis for U. Then*

$$R_K(\eta_1, ..., \eta_{r_1+r_2-1}) = [\mathcal{O}_K^*/\mu_K : U]R_K.$$

(2) *Let K be a totally complex number field of degree $2d$, and K^+ its maximal totally real subfield. Suppose $U \subset \mathcal{O}_{K^+}^*$ is a subgroup of the units, with a $\mathbb{Z}$-basis $\eta_1, ..., \eta_{d-1}$ given modulo ± 1. Then*

$$R_K(\eta_1, ..., \eta_{d-1}) = 2^{d-1} R_{K^+}(\eta_1, ..., \eta_{d-1}).$$

Proof. By the theory of elementary divisors, modulo torsion we may take a basis ϵ_i of fundamental units such that $\eta_i = \epsilon_i^{d_i}$, with $d_i \in \mathbb{Z}$, $\prod_i d_i = [\mathcal{O}_K^*/\mu_K : U]$. In this basis it is clear that

$$R_K(\eta_1, ..., \eta_{r_1+r_2-1}) = \det(\lambda_i \log|\tau_i(\eta_j)|) = \det(d_i\lambda_i \log|\tau_i(\epsilon_j)|) = [\mathcal{O}_K^*/\mu_K : U]R_K,$$

establishing part (1).

For part (2), note that since each of these units is real, R_K and R_{K^+} are computing the same expression except each λ_i is a 1 for K^+ and a 2 for K, so one sees a scaling by 2^d, as claimed. □

The lemma establishes that R_K/R_{K^+} is a rational number, and that to compute it we must compute $[\mathcal{O}_K^* : \mathcal{O}_{K^+}^*]$ (taking care to keep track of torsion).

Proposition 9.12. *Let $K = \mathbb{Q}(\zeta_n)$, and $\mu_n = (\mathcal{O}_K^*)^{tors}$. The index $Q := [\mathcal{O}_K^* : \mu_n\mathcal{O}_{K^+}^*]$ is equal to 1 if n is a prime power, and 2 otherwise.*

Proof. Firstly, let us show the index is 1 or 2. We define a homomorphism

$$\phi : \mathcal{O}_K^* \to \mu_K$$

by $\phi(u) = u/\bar{u}$, noting that this has absolute value 1 in all embeddings, so is a root of unity. We may compose ϕ with projection to μ_K/μ_K^2 to get a homomorphism into a subgroup of order 2. We claim the kernel of this map is exactly $\mu_K\mathcal{O}_{K^+}^*$. Indeed it visibly contains this group, but conversely suppose $\phi(u) = \eta^2$ for some $\eta \in \mu_K$. Then

$$\phi(\eta^{-1}u) = \eta^{-2}\phi(u) = 1$$

so $\eta^{-1}u \in \mathcal{O}_{K^+}$, as required.

Next, let us show that if $n = p^k$ is a power of an odd prime $p > 2$ then actually $\mathcal{O}_K^* = \mu_K\mathcal{O}_{K^+}^*$. We need to rule out the possibility that there is $\epsilon \in \mathcal{O}_K^*$ such that $\epsilon/\bar{\epsilon} = -\zeta^a$ for some a. But if $\epsilon = a_0 + a_1\zeta + ... + a_{(p-1)p^{k-1}-1}\zeta^{(p-1)p^{k-1}-1} \equiv a_0 + ... + a_{(p-1)p^{k-1}-1} \equiv \bar{\epsilon}$ mod $(1-\zeta)$, then

$$\epsilon = -\zeta^a\bar{\epsilon} \equiv -\epsilon \mod (1-\zeta)$$

so $(1-\zeta)|2\epsilon$. But since $p > 2$ and ϵ is a unit this is impossible.

Now, if $n = 2^k$ we need a different argument. Suppose $\epsilon/\bar{\epsilon} = \zeta$ a primitive 2^k-th root of unity. We can compute the norm $N_{\mathbb{Q}(\zeta)/\mathbb{Q}(i)}(\zeta) = \pm i$, via the factorisation

$$\frac{X^{2^k}-1}{X^{2^{k-1}}-1} = X^{2^{k-1}} + 1 = (X^{2^{k-2}} - i)(X^{2^{k-2}} + i).$$

But also $N(\epsilon)$ is a unit in $\mathbb{Q}(i)$ so must be one of $\pm i, \pm 1$. It's easy to see that none of these possibilities allow for

$$N(\epsilon)/\overline{N(\epsilon)} = \pm i.$$

Finally, suppose n has two distinct prime factors p, q, $\zeta = \zeta_n$ a primitive nth root of unity. Then we claim that looking at the expression

$$n = \prod_{i=1}^{n-1}(1-\zeta^i)$$

we can see that $(1-\zeta)$ is a unit. Indeed, for any $p^k||n$, we will have $\prod_{n/p^k|i}(1-\zeta^i) = u_p p^k$, for some unit u_p. Dividing through by this expression for all $p|n$, we see that a product of elements in $\mathcal{O}_K$ one of the factors of which is $(1-\zeta)$ is equal to a unit.

But $(1-\zeta)/(1-\zeta^{-1}) = -\zeta$, which we claim isn't in μ_n^2. Indeed, it's clear that $-\zeta$ generates μ_n, and $2||\mu_n|$ so $-\zeta$ can never be a square.

□

The upshot from this computation together with the previous lemma is the following.

Corollary 9.13. *Let $K = \mathbb{Q}(\zeta_n)$ and $Q = 1$ if n is a power of a prime, $Q = 2$ otherwise. Then*

$$\frac{R_K}{R_{K^+}} = \frac{2^{\phi(n)/2-1}}{Q}.$$

Proof. This is immediate from the lemma and the previous proposition. We get

$$R_K = Q^{-1} R_K(\mathcal{O}_{K^+}) = \frac{2^{\phi(n)/2-1}}{Q} R_{K^+}.$$

□

We turn to make some final remarks about the factor (h/h^+). A priori this is just some rational number, but in fact it is an integer with a definite interpretation.

Theorem 9.14. *Let $K = \mathbb{Q}(\zeta_n)$. The natural map $Cl(K^+) \to Cl(K)$ is injective. In particular $h^- := h/h^+$ is the order of the quotient $Cl(K)/Cl(K^+)$.*

Proof. Firstly let us remark what the natural map is. Given a class $C \in Cl(K^+)$ we take an ideal $\mathfrak{a} \in C$, and extend it to an ideal $\mathfrak{a}\mathcal{O}_K$ in $\mathcal{O}_K$, which has a class $[\mathfrak{a}\mathcal{O}_K] \in Cl(K)$. We need to check this map is well-defined. Given a different $\mathfrak{a}' \in C$, by the definition of ideal classes we can find $x, y \in \mathcal{O}_{K^+}$ such that $x\mathfrak{a} = x'\mathfrak{a}'$. But now it's obvious that

$$[\mathfrak{a}\mathcal{O}_K] = [x\mathfrak{a}\mathcal{O}_K] = [x'\mathfrak{a}'\mathcal{O}_K] = [\mathfrak{a}'\mathcal{O}_K].$$

It's now obvious this map is a group homomorphism.

To show it is injective (which is not true for general extensions), let us take $I \subset \mathcal{O}_{K^+}$ and suppose it becomes principal in $\mathcal{O}_K$. Let us suppose $I\mathcal{O}_K = (\alpha)$. We claim I is itself principal. Since I is a real ideal, we deduce that $(\alpha/\bar{\alpha}) = (1)$, so $\alpha/\bar{\alpha}$ is a unit with absolute value 1. This shows it is a root of unity.

Let us split into two cases. If n is not a prime power, $Q = 2$, and our analysis of units gives that there is some unit $\epsilon \in \mathcal{O}_K$ with

$$\epsilon/\overline{\epsilon} = \overline{\alpha}/\alpha.$$

But now $I\mathcal{O}_K = (\alpha) = (\alpha\epsilon)$ and $\alpha\epsilon$ is real. Since $I\mathcal{O}_K$ and $(\alpha\epsilon)$ have the same factorisation into primes in $\mathcal{O}_K$, we must have

$$I = (\alpha\epsilon) \subset \mathcal{O}_{K^+}$$

establishing the claim.

If $n = p^k$, take $\zeta = \zeta_{p^k}$ and let $\pi = \zeta - 1$ and note $\pi/\bar{\pi} = -\zeta$, which generates $\mu(\mathbb{Q}(\zeta))$. In particular this means that $\bar{\alpha}/\alpha = (\pi/\bar{\pi})^d$ for some d. Rearranging to

$$\bar{\pi}^d = \pi^d \alpha/\bar{\alpha}$$

and using the fact that $\pi^d\alpha$ and I are both real, and that real ideals acquire even π-adic valuation, we see that

$$d = v_\pi(\pi^d\alpha) - v_\pi(\bar{\alpha}) = v_\pi(\pi^d\alpha) - v_\pi(I) \in 2\mathbb{Z}.$$

In particular we see that $\bar{\alpha}/\alpha = \zeta'/\bar{\zeta}'$ for some root of unity ζ', so $\alpha\zeta'$ is real and $I = (\alpha\zeta')$ as before. □

After all this work, we see that our relative class number formula gives a simple relationship between h^- and the generalised Bernoulli numbers.

Theorem 9.15 (Relative class number formula, refined form)**.** *Let $K = \mathbb{Q}(\zeta_n)$, all other notation as introduced above. Then we have*

$$h^- = Q\omega_K \prod_{\chi \text{ odd}} (-\frac{1}{2}B_{1,\chi}).$$

It is striking that this is now a formula asserting an equality of two algebraic numbers, and in fact integers. Working away from the only potentially troublesome prime 2 we have the following immediate result.

Corollary 9.16 (Half of Kummer's Theorem)**.** *Let p be an odd prime. Suppose $p|B_m$ for some even m, $2 \leq m \leq p-3$. Then*

$$p||Cl(\mathbb{Q}(\zeta_p))|.$$

Proof. By the second Kummer congruence, $p|B_m$ implies $p|B_{1,\omega^{m-1}}$ and since m is even, ω^{m-1} is an odd character, so contributes to the right hand side of the above formula for $n = p$. Thus $p|h^-$ and in particular $p|h$. □

Let us remark that we have also reduced the other direction of Kummer's theorem to the statement that $p|h \Leftrightarrow p|h^-$. This is a theorem, which completes the proof of Kummer's result. It is worth remarking that Vandiver conjectured the following much stronger statement (which is still an open problem).

Conjecture 9.17 (Vandiver's Conjecture)**.** *Let $K = \mathbb{Q}(\zeta_p)$. Then p does not divide $h^+ = |Cl(\mathbb{Q}(\zeta_p + \zeta_p^{-1})|$.*

10. Gauss Sums and Stickelberger's Theorem

The class number formula gave us a relationship between a product of Bernoulli numbers and the order of a class group. However, the class group comes with an action of $Gal(K/\mathbb{Q})$ which can be used to break it into pieces, and to study these pieces we will need a "Galois-theoretic enhancement" of the class number formula. To this end we will define the *Stickelberger element* θ which is a formal linear combination of elements of $Gal(K/\mathbb{Q})$, and will play a role analogous to that of $L(1,\chi)$, and prove Stickelberger's theorem showing that (in a suitable sense) θ kills the class group (which is analogous to the class number formula).

Lemma 10.1 (Galois acts on class groups)**.** *Let $K/\mathbb{Q}$ be a Galois extension, $G = Gal(K/\mathbb{Q})$. Then G acts naturally on $Cl(K)$ via*

$$\sigma([\mathfrak{a}]) = [\sigma(\mathfrak{a})].$$

Proof. We need to check the above formula is well-defined. Suppose $x\mathfrak{a} = y\mathfrak{b}$. Then

$$\sigma(x)\sigma(\mathfrak{a}) = \sigma(x\mathfrak{a}) = \sigma(y\mathfrak{b}) = \sigma(y)\sigma(\mathfrak{b})$$

and we are done. □

At the heart of the proof of Stickelberger's theorem is the arithmetic of Gauss sums, so before we get there let's take the time to define and study them systematically (following Washington §6.1).

Let $\zeta_p \in \mathbb{C}$ be a fixed p-th root of unity (traditionally one takes $e^{2\pi i/p}$ but let's ignore this complex analytic ambiguity), and let κ be a finite field of order$q = p^d$. These data determine a canonical nontrivial additive character

$$(\psi, +) : \kappa \to \mathbb{C}^*$$

given by $\psi(a) = \zeta_p^{Tr_{\kappa/\mathbb{F}_p}(a)}$.

Now let $\chi : \kappa^* \to \mathbb{C}^*$ be a multiplicative character. For this section we will adopt the convention that $\chi(0) = 0$ even if χ is the trivial character (and in general these are no longer Dirichlet characters anyway unless $\kappa = \mathbb{F}_p$, so we should banish such thoughts!).

We can now define a *Gauss sum* to be the obvious Fourier-like thing

$$g(\chi) = -\sum_{a\in\kappa} \chi(a)\psi(a).$$

The sign convention is perhaps justified by the calculation

$$g(1) = -\sum_{a\in\kappa^*} \psi(a) = \psi(0) - \sum_{a\in\kappa} \psi(a) = 1.$$

We also remark that if $\chi^m = 1$, $g(\chi) \in \mathbb{Q}(\zeta_{pm})$, so one should really view such numbers as algebraic objects (sitting inside a field with a well-defined complex conjugation) even though one sometimes thinks of them as complex. It is also obvious from the definition that they are algebraic integers.

Lemma 10.2. *Let χ be a character. Then:*

(1)
$$g(\bar{\chi}) = \chi(-1)\overline{g(\chi)},$$

(2) *if $\chi \neq 1$,*
$$g(\chi)\overline{g(\chi)} = q,$$

(3) *if $\chi \neq 1$,*
$$g(\chi)g(\bar{\chi}) = \chi(-1)q.$$

Proof. We make calculations. For (1),

$$g(\bar{\chi}) = -\sum_{a\in\kappa} \bar{\chi}(a)\psi(a) = -\sum_{a\in\kappa} \overline{\chi(-1)\chi(-a)\psi(-a)} = \chi(-1)\overline{g(\chi)}.$$

For (2),

$$\begin{aligned} g(\chi)\overline{g(\chi)} &= \sum_{a,b\neq 0} \chi(ab^{-1})\psi(a-b) \\ &= \sum_{b,c\neq 0} \chi(c)\psi(bc-b) \\ &= \sum_{b\neq 0} \chi(1)\psi(0) + \sum_{c\neq 0,1} \chi(c) \sum_{b\neq 0} \psi(b(c-1)). \\ &= (q-1)+1 = q. \end{aligned}$$

The final equality is because when $c \neq 0, 1$, $\sum_{b\neq 0} \psi(b(c-1)) = -1$, so

$$\sum_{c\neq 0,1} \chi(c) \sum_{b\neq 0} \psi(b(c-1)) = -\sum_{c\neq 0,1} \chi(c) = 1 - \sum_{c} \chi(c) = 1.$$

Of course (3) follows directly from (1) and (2). □

Lemma 10.3. *Let χ_1, χ_2 be two characters of order dividing m. Then*

$$\frac{g(\chi_1)g(\chi_2)}{g(\chi_1\chi_2)}$$

is an algebraic integer in $\mathbb{Q}(\zeta_m)$.

Proof. If $\chi_1 = \chi_2^{-1}$, the previous lemma together with $g(1) = 1$ gives immediately that

$$g(\chi_1)g(\chi_2)/g(\chi_1\chi_2) = \begin{cases} 1 & \text{if } \chi_1 = \chi_2 = 1 \\ \chi_1(-1)q & \text{if } \chi_1 \neq 1. \end{cases}$$

If $\chi_1\chi_2 \neq 1$ we compute

$$\begin{aligned} g(\chi_1)g(\chi_2) &= \sum_{a,b} \chi_1(a)\chi_2(b)\psi(a+b) \\ &= \sum_{a,b;b\neq 0} \chi_1(a)\chi_2(b-a)\psi(b) \\ &= \sum_{b,c;b\neq 0} \chi_1(b)\chi_1(c)\chi_2(b)\chi_2(1-c)\psi(b) \\ &= g(\chi_1\chi_2)\sum_{c\in\kappa} \chi_1(c)\chi_2(1-c). \end{aligned}$$

□

Now we consider the action of $Gal(\mathbb{Q}(\zeta_{pm})/\mathbb{Q})$ on these Gauss sums. Let us assume $p \not| m$, so that $\mathbb{Q}(\zeta_m)$ and $\mathbb{Q}(\zeta_p)$ are linearly disjoint. Take $b \in \mathbb{Z}$ such that $(b, m) = 1$. We will denote by σ_b the Galois element

$$\sigma_b : \zeta_p \mapsto \zeta_p, \zeta_m \mapsto \zeta_m^b.$$

Lemma 10.4. *Suppose $\chi^m = 1$. Then*

$$g(\chi)^{b-\sigma_b} := \frac{g(\chi)^b}{g(\chi)^{\sigma_b}} \in \mathbb{Q}(\zeta_m)$$

and $g(\chi)^m \in \mathbb{Q}(\zeta_m)$.

Proof. Taking $b = m + 1$, in which case $\sigma_b = 1$ the second claim follows from the first. It suffices to check that for any $\tau \in Gal(\mathbb{Q}(\zeta_{mp})/\mathbb{Q}(\zeta_m))$, $(g(\chi)^{b-\sigma_b})^\tau = g(\chi)^{b-\sigma_b}$. Such τ is of the form $\zeta_p \mapsto \zeta_p^c, \zeta_m \mapsto \zeta_m$ for some $c \in \mathbb{Z}$. We may therefore compute

$$g(\chi)^\tau = -\sum \chi(a)\psi(ca) = \chi(c)^{-1}g(\chi)$$

and

$$(g(\chi)^{\sigma_b})^\tau = g(\chi^b)^\tau = \chi(c)^{-b}g(\chi^b).$$

Putting these together,

$$(g(\chi)^{b-\sigma_b})^\tau = \frac{(\chi(c)^{-1})^b}{\chi(c)^{-b}} g(\chi)^{b-\sigma_b} = g(\chi)^{b-\sigma_b}.$$

□

One more useful fact before we move onto Stickelberger.

Lemma 10.5. *Gauss sums are invariant under pth power:*

$$g(\chi^p) = g(\chi).$$

Proof. This is just a calculation, noting that $Tr(a) = Tr(a^p)$ since $a \mapsto a^p$ is an automorphism of κ fixing $\mathbb{F}_p$. We have

$$g(\chi^p) = -\sum_a \chi(a^p)\zeta_p^{Tr(a)} = -\sum_a \chi(a^p)\zeta_p^{Tr(a^p)} = g(\chi).$$

□

Now let us turn to Stickelberger's theorem. One has such a theorem for any abelian extension of $\mathbb{Q}$, but we will focus on $K = \mathbb{Q}(\zeta_m)$. Let $G = Gal(\mathbb{Q}(\zeta_m)/\mathbb{Q}) = (\mathbb{Z}/m\mathbb{Z})^*$.

We may form the *group algebra* $\mathbb{Z}[G] = \{\sum_{\sigma\in\mathbb{G}} n_i\sigma : n_i \in \mathbb{Z}\}$ with the obvious addition and multiplication. Since G is abelian, this is a commutative ring. It will also be convenient for us to work in $\mathbb{Q}[G] = \{\sum_{\sigma\in\mathbb{G}} n_i\sigma : n_i \in \mathbb{Q}\}$.

Recall by the first lemma of this section that $Cl(K)$ has an action of G. It is also an abelian group, and combining these two structures we see that $Cl(K)$ is a module over $\mathbb{Z}[G]$, and it will be our convention to write the action multiplicatively and on the right, so $\alpha \in \mathbb{Z}[G]$ will take $C \mapsto C^\alpha$.

Recall that $L(1, \chi)$ was related to $B_{1,\bar{\chi}}$ and (for $\chi \neq 1$ of conductor f) we have the formula

$$B_{1,\bar{\chi}} = \frac{1}{f} \sum_{1\leq a\leq f,(a,f)=1} a\chi(a^{-1}).$$

We wish to view this as the "evaluation at χ" of the *Stickelberger element*

$$\theta = \frac{1}{m} \sum_{a=1,(a,m)=1}^{m} a\sigma_a^{-1} \in \mathbb{Q}[G].$$

The main theorem (which one can see as a refinement of part of the statement of the class number formula) is the following.

Theorem 10.6 (Stickelberger's Theorem)**.** *Let $\beta \in \mathbb{Z}[G]$ be such that $\beta\theta \in \mathbb{Z}[G]$. Then for any ideal $\mathfrak{a} \subset \mathcal{O}_K$, $\mathfrak{a}^{\beta\theta}$ is principal. In other words, the ideal $I = \mathbb{Z}[G] \cap \theta\mathbb{Z}[G]$ annihilates the class group of K.*

References

[1] , .